DOCUMENT RESUME

ED 304 316 SE 050 373

AUTHOR Adams, Velma A., Ed.
TITLE Connecticut Assessment of Educational Progress 1984-85. Science Summary and Intrepretations.
INSTITUTION Connecticut State Dept. of Education, Hartford.
PUB DATE 86
NOTE 67p.; For earlier assessments, see ED 205 609-610.
PUB TYPE Reports - Descriptive (141) -- Reports - Evaluative/Feasibility (142) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Biological Sciences; Earth Science; *Elementary School Science; Elementary Secondary Education; Grade 4; Grade 8; Grade 11; Physical Sciences; *Sciences; *Secondary School Science; State Programs; *Testing Programs; Test Interpretation; *Test Results; Test Use
IDENTIFIERS *Connecticut Assessment of Educational Progress

ABSTRACT

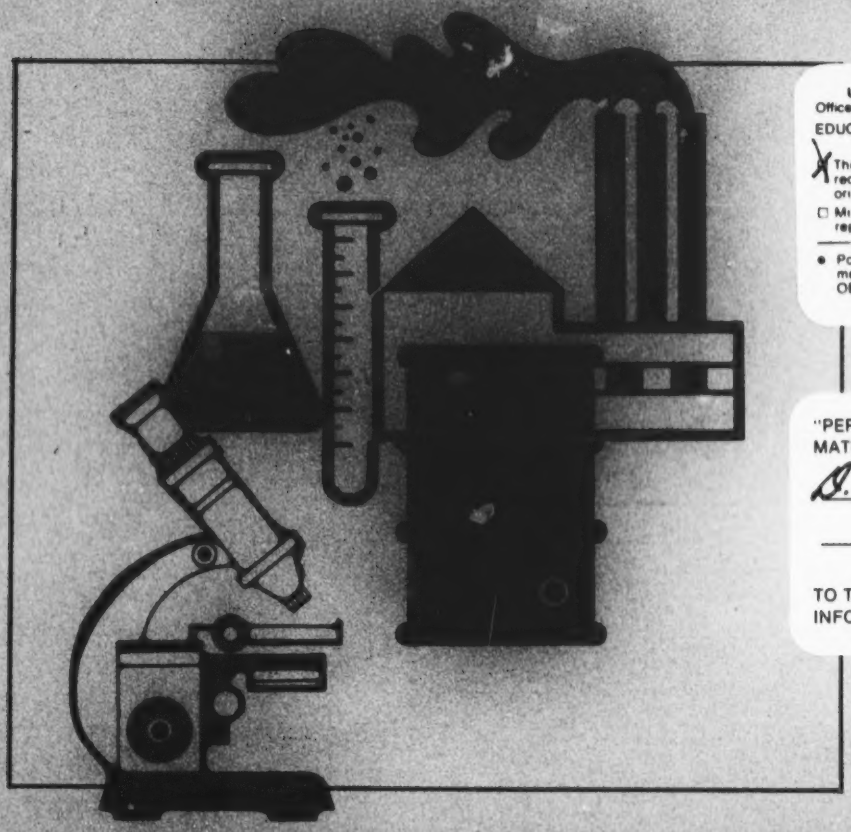
The Connecticut Assessment of Educational Progress (CAEP) is modeled after the National Assessment of Educational Progress (NAEP) and has been in operation since 1971. This report is the culmination of a multi-year effort by the Connecticut State Department of Education to make its CAEP program even more useful to classroom teachers and other local district personnel in improving curriculum than it has been in the past. The 1984-85 Science Assessment was designed to determine what Connecticut students know about topics in life science, earth science, and physical science; how they are able to use the processes of science such as measurement, observation, making inferences, and designing experiments; and how they feel about their science experiences and the methods used to teach them. To highlight the importance of this instructional approach and to determine how well Connecticut students perform with various types of equipment necessary for learning science, a practical test was also administered. Grades 4, 8, and 11 were tested and those test results and interpretations are provided for four catagories: scientific inquiry; life sciences; earth and space sciences; and the physical sciences. Comparisons are provided for statewide, national, and regional results. A sample test is included. (MVL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

CONNECTICUT ASSESSMENT OF EDUCATIONAL PROGRESS 1984-85

ED304316

SCIENCE



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Headspith

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SUMMARY AND INTERPRETATIONS

STATE OF CONNECTICUT DEPARTMENT OF EDUCATION - 1986

Connecticut State Board of Education

Abraham Glassman, Chairman
James J. Szerejko, Vice Chairman
A. Walter Esdaile
Warren J. Foley
Dorothy C. Goodwin
Rita L. Hendel
John F. Mannix
Julia S. Rankin
Humberto Solano

Norma Foreman Glasgow (ex officio)
Commissioner of Higher Education

Gerald N. Tirozzi
Commissioner of Education

Frank A. Altieri
Deputy Commissioner
Finance and Operations.

Lorraine M. Aronson
Deputy Commissioner
Program and Support Services

CONNECTICUT ASSESSMENT

OF

EDUCATIONAL PROGRESS

1984-85

SCIENCE

SUMMARY AND INTERPRETATIONS

For further information regarding the assessment, contact Joan Boykoff Baron (203-566-3847) or Sigmund Abeles (203-566-4825) at the Connecticut State Department of Education, Box 2219, Hartford, CT 06145.

CONTENTS

LIST OF TABLES AND EXHIBITS	iv
FOREWORD	v
PREFACE	vii
ACKNOWLEDGMENTS	ix
CHAPTER I: INTRODUCTION	1
Overview	1
Instruments	1
Procedures	3
CHAPTER II: COMPARATIVE ACHIEVEMENT RESULTS	6
Introduction	6
Comparisons to Previous Statewide, National and Regional Results	6
Performance by Gender and Type of Community	7
CHAPTER III: ACHIEVEMENT RESULTS BY SCIENCE DISCIPLINE	9
Interpreting Item-Level Results	9
Scientific Inquiry	9
Life Sciences	17
Earth and Space Sciences	22
Physical Sciences	26
CHAPTER IV: RESULTS OF THE PRACTICAL TESTS	31
Grade 4 Results	31
Grade 8 and 11 Results	32
CHAPTER V: QUESTIONNAIRE RESULTS	37
Student Questionnaire	37
Teacher Questionnaire	40
Principal Questionnaire	44
CHAPTER VI: SUMMARY AND RECOMMENDATIONS	46
Comparisons with the Nation, the Northeast and Connecticut Five Years Ago	46
Connecticut Strengths and Weaknesses	47
Selected Questionnaire Results	49
Recommendations	51
APPENDIX: SAMPLE TEST	53

TABLES

1. Distribution of Items by Content Categories 2
2. Students and Schools Participating in the Assessment 4
3. Average Deviations in Percent Correct between 1984-85 Connecticut Results and Previous Connecticut, U.S., and Northeast Results 7
4. Achievement Results by Science Area by Gender and by Type of Community 8
5. Results on Selected Human Biology Items 19
6. Results on Selected Multiple-Choice Items about Pollution 22
7. Results on Selected Multiple-Choice Items about the Nature/Structure of Matter 29
8. Students' Experience with Laboratory Equipment 34
9. Previous Experience and Performance with Triple-Beam Balance 34
10. Approaches Used to Solve Survival Problem 35
11. Control of Variables--Performance Test vs. Multiple-Choice Test 36
12. Frequencies of Selected Classroom Activities and Their Relationships to Achievement 39
13. Science Coursework of Grade 11 Students 40
14. Types of Questions Teachers Ask 42
15. Factors Most Frequently Cited as Serious Problems 43

EXHIBITS

1. Results on Selected Items about Awareness of Science 10
2. Test Question on Sequence of Events 13
3. Applying Measurement Skills 15
4. Interpretation of a Table--Grade 4 16
5. Understanding of Weather Map and Relative Humidity 24
6. Results on Selected Items in Mechanics 27
7. Paper Helicopter 32
8. The Survival Exercise 34

The Connecticut Assessment of Educational Progress (CAEP) is modeled after the National Assessment of Education Progress (NAEP) and has been in operation since 1971. Statewide assessments have been conducted in art, career education, citizenship/social studies, mathematics, music, reading and science. The 1984-85 Science Assessment is the third science assessment sponsored by the Connecticut State Department of Education. It was conducted by Advanced Systems in Measurement and Evaluation of New Hampshire.

The 1984-85 Science Assessment was designed to determine what Connecticut students know about topics in life science, earth science and physical science; how they are able to use the processes of science such as measurement, observation, making inferences and designing experiments; and how they feel about their science experiences and the methods used to teach them.

An important component of this assessment was a practical test. Science, by its very nature, is a "doing" activity. To highlight the importance of this instructional approach and to determine how well Connecticut students perform with various types of equipment necessary for learning science, a practical test was administered. Its purpose was to assess students' skills with laboratory devices and their ability to solve problems involving the use of these devices. In a number of instances, the questions on the pencil and paper instrument addressed the same problems as were given on the practical test. This technique provided information on the adequacy of each type of instrument to assess student achievement and ability in problem solving.

This year 10,611 students in grades 4,8, and 11 were tested in the statewide science sample. Another 8,558 students were added as part of the local option portion of this program. The 144 local schools which elected to participate in the local option administered the same tests to their students and have received achievement results for individual students, schools and the district.

I want to thank all the students, teachers and administrators throughout the state who made this assessment possible. The cooperation, time and effort of all participants are greatly appreciated. I am confident that teachers of science and curriculum specialists will find this report useful in their ongoing efforts to strengthen programs in Connecticut schools.

Gerald N. Tirozzi
Commissioner of Education

This report is the culmination of a multi-year effort by the Connecticut State Department of Education to make its CAEP program even more useful to classroom teachers and other local district personnel in improving curriculum than it has been in the past. The individually administered practical exercises and the improved district-level reports for districts participating in the local option testing are reflective of this effort. We believe that teachers and curriculum leaders will discover that many of the findings provided in this report have practical significance to them as they work to improve curriculum and instruction in science.

A pervasive problem inherent in statewide assessments is the tendency for the staff in any particular school district to accept the results as true elsewhere, but not in their own schools. Whereas weaknesses found statewide do not apply to every district, it is likely that even the best-performing districts will have some weaknesses that are similar to the ones uncovered statewide. Thus, we have included in an appendix to this report a sample test consisting of selected items from the assessment along with the statewide results for each of the selected items. Teachers may copy the sample test, administer it to their students, and determine whether certain problem areas identified statewide hold true for their classes as well. If not, they can justifiably say that the weaknesses identified in statewide trends are not present locally. However, if their students score poorly on these questions, they should read with extra care the results and recommendations in this report.

Approximately 300 test items were used in this assessment; of them, only a handful are reproduced in this report. Readers wishing more detail about the assessment instruments and results are referred to the Technical Report for this assessment, available through the Connecticut State Department of Education.

ACKNOWLEDGMENTS

The 1984-85 Connecticut Assessment of Educational Progress (CAEP) in science was a joint effort of the Office of Research and Evaluation and the Bureau of Curriculum and Staff Development at the Connecticut State Department of Education (CSDE), and Advanced Systems in Measurement and Evaluation, Inc. of Dover, New Hampshire.

The assessment involved substantial effort on the part of many persons. Joan Baron, Sigmund Abeles, and Douglas Rindone of the Connecticut State Department of Education, were responsible for the overall direction of the project. Substantial contributions were also made by the Statewide Advisory Committee for Science (shown below). At Advanced Systems, the project was implemented by Stuart Kahl and Richard Hill.

Statewide Advisory Committee

George C. Atamian	Bausch and Lomb Corporation, East Hartford
Eleanor Bennett	Ridge Road Elementary School, North Haven
Ronald Bergeron	Brooklyn School
Dr. James Della Valle	Northeast Utilities, Hartford
Victor Giangraves	Ridgefield High School
Delilah Gomes	West Haven High School
Richard Mace	Bridgeport Public Schools
Louise Mastracchio	Bethel Middle School
Douglas Murray	Chaplin Elementary School
June A. Rackliff	Bozrah Fields Memorial School
Dr. Frank Salamon	St. Joseph College, West Hartford
Mary Lou Sette	Roxbury School, Stamford
Maureen Smith	formerly of Jonathan Law High School, Milford
Wilma Toney	Weaver High School, Hartford
Dr. Charles Vasquez	Darien High School
Dolores Vecchiarelli	Long Lots Junior High School, Westport
Joseph Walczak	H.C. Wilcox Technical School, Meriden
Dr. Ralph Yulo	Eastern CT State University, Willimantic

The following persons were responsible for test coordination and test administration in conjunction with the practical testing which involved the individual testing of approximately 900 students across the state:

Marguerite Black, Wilton	Thomas Jankowic, East Norwalk
Mark Bosworth, Enfield	Norma LoMonaco, West Willimantic
Lee Colber, Norwalk	Vicky Nietch, Torrington
Jeffrey Croce, Waterbury	Patricia O'Neil, Waterbury
Carolyn Dahl, Middlebury	Muriel Sandquist, New Britain
Joel Greene, Westport	Constance T. Shea, Windsor
Marshall Hahn, East Haven	Lawrence Shea, Branford

CHAPTER I

INTRODUCTION

Overview

The Connecticut Assessment of Educational Progress (CAEP) is a program established by the Connecticut State Department of Education (CSDE) to provide for the periodic monitoring of statewide performance of students in grades 4, 8, and 11 in curriculum areas selected for assessment. Tests designed specifically for CAEP are administered to a sample of students statewide, and the results are disseminated in a variety of ways.

The goals of the CAEP assessments are to:

- provide achievement results useful for decision making regarding curricula and instruction at both the state and local levels;
- collect baseline and longitudinal data for determining student growth in knowledge and skills in subsequent years;
- compare the achievement of Connecticut students with students nationally when possible; and
- encourage local school districts to utilize CAEP instruments and procedures for local evaluation and planning.

Since the program began in 1971, CAEP has assessed reading/language arts, computer literacy, mathematics, citizenship/social studies, science, career education, art and music, and business and office education. The 1984-85 assessment was the third one in the area of science.

The science assessment was conducted by Advanced Systems in Measurement and Evaluation, Inc. of Dover, New Hampshire, under contract to the Connecticut State Department of Education. Staffs of both Advanced Systems and CSDE worked jointly on all aspects of the assessment program, including instrument development, test coordination, and the analysis and reporting of results. Additional support was provided by an advisory committee of educators and representatives of the science/business community in the state of Connecticut. (See "Acknowledgments.")

Instruments

Group-Administered Tests

Two nonequivalent test forms were administered at each of grades 4, 8, and 11. Each student in the statewide sample completed one of these forms. Most of the test items were multiple-choice; a few items in each form, however, were short answer questions requiring hand scoring. Table 1 shows the distribution of test items across content categories for each of the three grade levels tested.

Table 1

Distribution of Items by Content Categories

Category	Number of Test Items Grade 4	Grade 8	Grade 11
Scientific Inquiry			
Awareness of Science and Scientific Processes	3	4	9
Designing Experiments	3	5	5
Observing and Measuring	8	4	5
Interpreting and Translating Data	7	3	6
Drawing Conclusions and Inferences	4	5	4
Life Sciences			
Characteristics of Life	4	5	5
Animal Life	8	14	13
Plant Life	10	5	3
Ecology and Environment	4	8	9
Earth and Space Sciences			
Astronomy	7	7	7
Climate and Weather	5	5	7
Geology	6	7	9
Oceanography	2	5	3
Physical Sciences			
Physics	15	19	19
Chemistry	12	14	16
TOTAL NUMBER OF ITEMS	98	110	120

In addition to the above content categories, another dimension was used to classify test items. Every item was assigned to a cognitive process level--measuring knowledge, comprehension, application, or the higher processes of analysis, synthesis, and evaluation. These levels, representing levels of Bloom's Taxonomy, have been used by the National Assessment of Educational Progress. Their use assures that the test questions as a set do not overemphasize one particular level such as knowledge (recall) at the expense of other important cognitive processes.

Current literature on scientific literacy reflects two concerns:

- concern for the needs of the general public in terms of the knowledge needed to function effectively as an individual and as a citizen in a modern, technological society; and

- concern for the adequacy of training of persons entering scientific or technological fields.

Therefore, particular attention was paid during test development to assure that the tests included items addressing knowledge and skills needed by all students

as well as items more appropriate for students who are likely to pursue science academically or professionally.

Practical exercises requiring scientific apparatus were administered to a subset of the students in the statewide sample. These exercises, given individually, asked the students to conduct simple investigations, use measuring devices, and design experiments.

To provide a more complete picture of the status of science education in Connecticut, students, teachers and principals responded to questionnaires addressing matters such as student and school staff backgrounds, attitudes, and experiences; program characteristics; materials and equipment used. Emphasis was placed on variables expected to relate to student achievement which are "alterable." Thus, this information was intended to augment the information on strengths and weaknesses obtained from the achievement data and to be of use in guiding program improvement.

Procedures

Instrument Development

Between April and December of 1984, staff members of the Connecticut State Department of Education and Advanced Systems met several times with the State Advisory Committee. At the earliest of these meetings, objectives were decided upon. Advanced Systems staff then prepared tentative item sets using a variety of sources including previous Connecticut, national, and other assessments. New items were developed to meet the coverage of objectives identified by the advisory committee. At later committee meetings, items for the various assessment instruments were reviewed and then approved, modified, or rejected.

Pilot testing was conducted to assure the appropriateness of the test items. Timing information and item difficulty levels were obtained, so that the overall difficulty of the instruments would be kept at an appropriate level and so that the difficulty of the two test forms at each grade level could be balanced. Additional information regarding the clarity and complexity of items was obtained from interviews with students.

After final instruments were constructed by Advanced Systems and reviewed by CSDE, they were typeset and printed by Wenday, Inc. of Concord, New Hampshire. Machine-scorable answer sheets, tailored to the tests, were designed by Cooperative Educational Services of Norwalk, Connecticut and produced by National Computer Systems of Lancaster, Pennsylvania.

Coordination

Continuing the practice begun in a previous CAEP assessment, Advanced Systems drew a statewide sample of schools that excluded those that had participated in the previous year's sample assessment. Somewhat fewer schools were involved in the study but the number of classes sampled in each was increased so that meaningful results could be made available to schools in the sample without the district's having to "buy" the local option. As in previous years, several districts elected to participate in the local option where they tested many, if not all, of their students at appropriate grade levels to

obtain detailed reports on the science achievement of their students. Table 2 shows the extent of the participation in the science assessment.

Table 2

Students and Schools Participating in the Assessment

	Grade 4	Grade 8	Grade 11
Number of Schools			
Statewide Sample Only	52	43	41
Statewide Sample and Local Option	21	14	14
Local Option Only	49	23	23
Total	122	80	78
Number of Students			
Statewide Sample Only	2053	1998	1642
Statewide Sample and Local Option	643	2075	2200
Local Option Only	2542	2920	3096
Total	5238	6993	6938
Number of Test Forms Administered			
Statewide Sample Only	2053	1998	1642
Statewide Sample and Local Option	728	3179	3248
Local Option Only	3230	3668	4723
Total	6011	8845	9615

In the fall of 1985, the superintendents were notified about the schools in their districts selected for the statewide sample and asked to submit orders if they elected to participate in the local option. Approximately two months before testing, principals of sampled schools were asked to provide information on the class groupings so that Advanced Systems could identify the classes to be tested. One month prior to testing, the test coordinators named by the superintendents and principals were sent confirmation of the quantities of testing materials to be shipped, classes to be tested, and other pertinent information. Materials were shipped the week prior to testing.

Administration

The bulk of the testing took place during two-week intervals in November, February, and April at grades 4, 8, and 11 respectively. The first week of each testing period was used for scheduled group testing and the second week was reserved for make-up testing. Some districts in the local option chose to test approximately four weeks later utilizing used test booklets, thereby saving on expenses. The group tests were administered by local school personnel following detailed instructions provided in an instruction manual.

Practical Tests

A subsample of schools (thirty at each grade level) was identified for participation in the practical testing of individual students. Exercises requiring students to measure things, manipulate objects, conduct simple investigations, etc. were given to ten students in each of these schools by external, trained administrators. This part of the assessment was conducted after the group testing and was coordinated entirely by the trained administrators. These administrators contacted the schools in advance and arranged for a testing room and for the schools to supply just a few of the required materials (hot water, microscope, balance). The student answer sheets completed during the group testing were used by the administrators to select the students for the practical tests--i.e., the administrators randomly selected answer sheets from the stack of completed answer sheets in each school. The students were sent to the special testing room two at a time. The two students being tested at one time worked independently on completely different exercises and were observed at the appropriate times by the administrator. The practice of testing two students at a time had been used in other programs to reduce examinee anxiety. After each pair of students finished, the administrator encoded their responses to the various exercises on the machine-scorable answer sheets using detailed scoring guides for the exercises.

Scoring, Analysis, Reporting

A few of the items in the group-administered tests were not multiple-choice and required students to write answers on a portion of the answer sheet designed for that purpose. Students' responses to these items were scored at Advanced Systems. The scorers used scoring guides to determine response categories which were encoded on the students' answer sheets. Once all hand scoring was completed, the answer sheets were shipped to Cooperative Educational Services (CES), where they were optically scanned. CES also produced a raw data tape. All computer analyses, including the production of the local option reports, were conducted by Advanced Systems. The statewide results are summarized in subsequent chapters of this document and reported in greater detail in the Technical Report for this assessment. Interpretations, conclusions, and recommendations presented in this report represent the joint efforts of the CSDE and Advanced Systems staffs and the science advisory committee.

CHAPTER II

COMPARATIVE ACHIEVEMENT RESULTS

Introduction

In this chapter, the 1984-85 performance levels in science of Connecticut students are compared to national and northeast performance levels based on results for items used in both the National Assessment of Educational Progress (NAEP) and the CAEP science assessment. Changes in performance in Connecticut between 1979-80 and 1984-85 are reported as well, since a number of the items used in this year's assessment were administered in the previous CAEP assessment of science. The levels of performance of various subgroups of Connecticut students are also compared, based on gender and type of community.

All assessments like CAEP are based on the testing of samples of students, and are subject to random errors created by sampling fluctuations. To avoid overinterpretation of small differences, the reader is advised to assume that, in general, **differences of fewer than four percentage points should be considered inconsequential.** In addition to sampling error, other factors must be considered when interpreting comparative data. These factors vary depending on the comparisons being made and will be discussed in the appropriate sections.

In this chapter, performance levels on the total test and subsets of the test items are compared. Chapter III makes national, regional, and grade level comparisons on individual items and evaluates the absolute performance levels on the test items. The discussion of student questionnaire results in Chapter V compares the total test performance of many groups defined by questionnaire item responses.

Comparisons to Previous Statewide, National, and Regional Results

While it is tempting to believe that direct comparisons to national norms can be obtained from the administration of sets of items common to statewide and national assessments, there are so many important differences in procedures that any results must be viewed with extreme caution. One is the problem of the year in which the national data were collected. Most of the science items for which data were available were administered nationally in 1976-77, although a few items date from 1981-82. Furthermore, NAEP sends teams of trained administrators to the schools where the testing is to be done. In the past, when testing began, the administrator started a special, paced audiotape which, in effect, read all directions, test questions, and answer options to the students as they read along in the test booklets. Practices used in the State of Connecticut differ considerably from many employed by NAEP. Those used by NAEP in the past are simply too costly for most states to follow in their own testing programs. As a consequence, comparisons of students in Connecticut, as measured by CAEP, to students nationwide or regionally, as measured by NAEP, must be made with caution.

The comparisons reported in this section are based on the most compatible data that can be obtained statistically. For example, the previous NAEP and

CAEP tested thirteen-year-olds while this year's CAEP tested eighth graders. For purposes of comparison, only data based on thirteen-year-olds in the eighth grades from any of the assessments were used. The same approach to age and grade eligibility was used for the grade 4 and 11 comparative results as well. Table 3 shows average deviations in percent correct between 1984-85 Connecticut results and U.S., Northeast, and previous Connecticut results for various subsets of test items.

Table 3

Average Deviations in Percent Correct Between 1984-85 Connecticut Results and Previous Connecticut, U.S., and Northeast Results

(n = number of items)

Topic	Grade 4			Grade 8			Grade 11		
	CT	US	NE	CT	US	NE	CT	US	NE
Scientific Inquiry	+3.5 n=6	+7.7 n=7	+4.7 n=7	+0.3 n=4	+5.8 n=4	+7.0 n=2	-5.0 n=4	-0.6 n=9	-4.0 n=9
Life Sciences	+2.0 n=4	+0.9 n=8	-0.1 n=8	-2.0 n=9	-0.3 n=11	-4.3 n=6	-0.1 n=8	0.0 n=10	-4.3 n=10
Earth and Space Sciences	+2.8 n=8	+5.1 n=7	+4.6 n=7	+0.8 n=4	+2.0 n=8	+0.3 n=6	-1.7 n=6	0.6 n=9	-1.2 n=9
Physical Sciences	+3.0 n=8	+1.0 n=11	-0.5 n=11	-2.0 n=10	+2.2 n=13	-1.5 n=10	+0.9 n=13	+2.1 n=10	+0.1 n=10
Total Test	+2.9 n=26	+3.3 n=33	+1.8 n=33	-1.3 n=27	+1.8 n=36	-1.0 n=24	-0.6 n=31	+0.3 n=38	-2.3 n=38

The most striking feature in the data above is the great consistency among previous national, regional, and statewide results. At grade 4, Connecticut students do show signs of strength relative to their national and regional counterparts in the areas of Scientific Inquiry and Earth/Space Sciences. Based on fewer items, the grade 8 students appeared to have an edge in Scientific Inquiry. As for change over time, only the grade 4 students showed a small, but consistent, improvement since the previous statewide assessment of science. On the average, grade 8 and grade 11 performance changed very little.

Performance by Gender and Type of Community

Table 4 shows the average science test scores for males and females as well as for students from various types of communities. Again, certain cautions should be kept in mind when interpreting these findings. First, it should be remembered that small differences in achievement scores between groups are not likely to be educationally meaningful. Second, it should be noted that observed differences do not indicate cause-effect relationships. While there may be a relationship between achievement and a given variable (e.g., gender of student), the data do not indicate the reason or cause for the relationship.

All districts in Connecticut are included in one of six types of community, based on U.S. Census Bureau data. These TOCs include large cities (towns with population of more than 100,000), fringe cities (towns contiguous with large cities and with populations of over 10,000), medium cities (towns with populations between 25,000 and 100,000 and not fringe cities), and small towns (towns with populations less than 25,000) divided into suburban, emerging suburban, and rural communities.

Table 4

Achievement Results by Science Area by Gender and by Type of Community

Average Percentage of Questions Answered Correctly

Science Area	Grade 4		Grade 8		Grade 11	
	Male	Female	Male	Female	Male	Female
Scientific Inquiry	56	55	56	58	53	55
Life Sciences	56	54	52	50	52	52
Earth/Space Sciences	52	48	47	43	52	45
Physical Sciences	56	51	54	47	54	46
Total Science	55	52	52	49	53	49

Type of Community	Grade 4	Grade 8	Grade 11
Large Cities	42	41	38
Fringe Cities	58	51	52
Medium Cities	51	50	51
Small Towns	58	55	55
Statewide Average	53	51	51

Overall, males slightly outperformed females at all three grade levels, but these differences should not be considered educationally significant. In scientific inquiry and life sciences, male and female performance levels were quite similar. However, in earth/space sciences and physical sciences, males had a sizable edge, particularly at the higher grades. The grade 11 differences might be explained in part by differences in the course-taking patterns of males and females. (See questionnaire results on page 40.)

Students in "Big Cities" scored well below statewide averages at all three grade levels. Students in "Small Towns" scored somewhat higher than students statewide, and students in "Fringe" or "Medium Cities" generally scored very near the statewide averages. The exception to the latter was the grade 4 students in "Fringe Cities" who performed approximately the same as the grade 4 students in "Small Towns." For the most part, the findings in Table 4 are consistent with the results of other Connecticut testing programs.

CHAPTER III

ACHIEVEMENT RESULTS BY SCIENCE DISCIPLINE

Interpreting Item-Level Results

The CAEP assessment was a survey of statewide achievement in science. As such, its instruments were developed to cover as broad a range of concepts and topics as possible rather than a selected few "essentials." While a few concepts were addressed by several items, this was not typically the case. The general character of assessment instruments, then, makes the discussion of individual item results all the more important.

One can have a great deal of faith in our ability to generalize from these results to how a larger population of students might have performed on these items because of the size and quality of the CAEP sample. However, when one attempts to generalize about performance in subdomains of science on the basis of a few items, one must be more cautious. Nevertheless, clear patterns do emerge from the item results. These patterns and supporting evidence are discussed in this section of the report. The discussion is organized by the science content outline shown on page 2.

This section is really a summary report. The reader is encouraged to review in greater detail the data provided in the Technical Report prepared for the CAEP assessment. Attention should be given not only to percentages of correct responses (p-values) but also to the percentages of students selecting particular incorrect responses. As later discussions will illustrate, item results may not be disturbing when only the p-values are considered, but could lead to far greater concern when misconceptions held by large numbers of students are discovered through examining the percentages of students selecting particular wrong responses.

Average p-values for the different subtopics within science are not reported. It would be inappropriate to claim that performance in one area is better than performance in another from such statistics. The items in one area might simply be more difficult. Performance on different sets of items cannot be compared in that manner. However, the substantive discussion of specific items, together with subjective but informed interpretation, does uncover strengths and weaknesses in the performance of Connecticut students.

Scientific Inquiry

Awareness of Science/Science Processes

Several CAEP items dealt with the awareness of the nature of science and scientific processes. A long-time concern of science educators has been the misconceptions and stereotypes students hold regarding science--beliefs which make science seem sterile and elitist and which may cause many students to believe that science is not an area for them to pursue. The results on three of the items shown in Exhibit 1 suggest that there is still reason for concern.

Item A suggests that many fourth graders associate science with conducting experiments and fail to see the roles of observation and measurement in

science. On Item B, many fourth graders showed that they believed scientists work only with hard facts and require laboratories. On the corresponding item for eighth and eleventh graders (Item C), there was some improvement, but still there was evidence that many students hold the same misconceptions at the higher grade levels.

Exhibit 1

Results on Selected Items about Awareness of Science

Percent Correct Gr. 4		
		Item A: Which of the following could a scientist do to find out how one rock is different from another?
17		A. observe
14		B. measure
31		C. experiment
38		*D. all of the above
1		no response
		Item B: Which statement about scientists is TRUE?
44		*A. When scientists solve one problem, they often find new problems.
26		B. Scientists work only with facts, not guesses.
7		C. Scientists know everything about what makes up things.
24		D. To do an experiment, a scientist needs a laboratory.
1		no response
Gr.8	Gr.11	Item C: Which of the following statements about science or scientists is generally true?
58	73	*A. The solution a scientist finds to one problem often raises new questions or problems.
14	10	B. Scientists deal with facts and rules and not with assumptions and uncertainty.
12	6	C. Scientists know all there is to know about many areas such as cell reproduction and the composition of the earth.
15	10	D. True scientific experiments must be conducted in laboratories for one to be sure of the results.
		Item D: Based on data from an experiment with plants, which statement below might be an <u>inference</u> rather than an <u>observation</u>?
36	50	*A. Lack of light causes plants to grow at a slower rate.
25	20	B. All of the plants are bent toward the light source
19	14	C. Each of the plants exhibits new growth at its tip.
18	16	D. Several of the plants have yellow leaves.
1	1	no response

*correct response

The grade 8 and 11 performance on Item D (also included above) was quite disappointing considering the skill measured is addressed repeatedly in reading and social studies as well as in science classes. Distinguishing between inferences and observations (or related concepts) is a critical thinking skill currently of great concern to educators.

Grade 11 students appear to be confused about scientific models, theories, and hypotheses. More than half of these students responded that a scientific model is "a small copy of a larger, more complicated piece of equipment," instead of the correct response, "a scheme that guides one's thinking about something not completely known." Only 28 percent of the eighth graders and 30 percent of the eleventh graders knew that a scientific theory is "a useful explanation of something which can be changed because of new evidence." Forty-one percent of the students at grade 8 and 47 percent of the students at grade 11 indicated that a theory is "a hypothesis that has been proven as a result of an experiment." Grade 8 and grade 11 performance was generally good on items addressing the need for scientists to be concerned about possible harmful side effects associated with scientific developments or environmental legislation.

Design of Experiments

Two general conclusions can be drawn from the results on items pertaining to the design of experiments. First, students show great faith in the written word in situations in which empirical verification is warranted. Two-fifths of the fourth graders and two-thirds of the eleventh graders recognized that measuring the height of all boys and girls in a school was the surest way to find out whether boys in the school are taller than girls in the school on the average. Interestingly, large percentages at both grade levels opted for looking up average heights in an encyclopedia or in the school records showing heights of students upon entering school. On another item, 42 percent of the fourth graders recognized that lifting heavier and heavier weights until a string broke is the best way to determine the strength of a string. Yet another 30 percent answered, "find out what the string is made of." Performance on this latter item was reasonably good, but it does reveal a tendency to rely on nonempirical approaches.

The second major conclusion is drawn from results on items concerning experimental controls. Students apparently are used to dealing with one factor at a time and have difficullty with the idea that situations may call for the control of several variables at once. One free-response question required eleventh graders to design an experiment comparing the effectiveness of two laundry detergents. Variables considered in the scoring of responses were type of dirt, type of fabric, amount of detergent, and washing procedures (water level, temperature, etc.). Only 14 percent of the students designed experiments that controlled three or four of these variables. Another 36 percent controlled one or two of the variables. Amount of detergent and washing procedures were the variables most ignored. Performance was not appreciably better on items asking students how to improve an experiment which included only one uncontrolled variable. Only 24 percent of the grade 8 students and 46 percent of the grade 11 students recognized the problem in an experiment about plant growth in which the amount of light was not controlled. On a simpler, but similar question, 61 percent of the grade 4 students

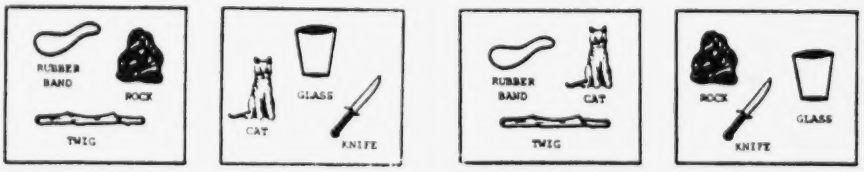
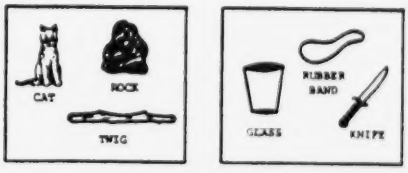
recognized that light needed to be controlled. On the practical tests described in a later section, student performance suggests that there may be less of a problem here than meets the eye. In actually conducting experiments, there are times when students encounter the variables requiring control one at a time, and in these instances they do tend to control such variables.

Observing and Measuring

There were no surprises on questions requiring students to observe. Most fourth graders (92 percent) correctly identified a set of pictures showing how an ice cream cone changes as it is eaten, shown below.

Percent Correct Gr. 4		
2	A.	
92	*B.	
2	C.	
1	D.	
1	E.	I don't know.

On a more difficult item, 55 percent of the grade 4 students selected the pair of pictures in which six objects were sorted into natural and man-made things. The objects were a rubber band, a glass, a knife, a cat, a rock, and a twig.

17	A.	 A. B.
10	B.	
55	*C.	 C.
17	D.	I don't know.
2		no response

The item shown below was administered at grades 4 and 8.

Exhibit 2.

Test Question on Sequence of Events

The pictures below were taken at the same location, but at different times. Which picture was taken LAST? (For eighth graders, "which picture was taken THIRD?")

Percent Correct Gr. 4		
32	A.	
3	B.	
3	C.	
60	*D.	
2	no response	

Gr. 8		
61	*A.	
3	B.	
4	C.	
31	D.	
1	no response	

Performance was almost identical at both grade levels (60 percent at grade 4 and 61 percent at grade 8). Although the wording of the item in the grade 8 test would make the question a little more difficult, one would still expect a much greater difference in the percents correct, based simply on age, education, and experience differences between the grades. Ninety-two percent of the fourth and eighth graders narrowed the choices down to the two pictures with the tree fallen--the most obvious difference in the pictures. Apparently, the older students could not examine the pictures for additional detail any better than the younger ones.

Forty-seven percent of the eleventh graders identified the proper sequence of the pictured stages of mitosis. While many students may not be familiar with mitosis, the item stated that this was a process of cell division. Too many students selected answer options that were unreasonable -- for example, options which did not show a simple single cell first or two distinct cells last.

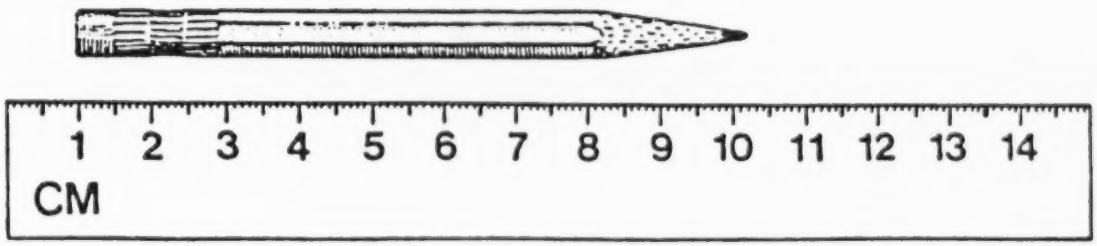
Students at all grade levels apparently comprehend some very basic measurement concepts, but have difficulty with more subtle concepts and with actual applications of measurement skills. Just under two-thirds of the fourth graders knew that a thermometer (as opposed to a battery, magnifying glass, or test tube) is most often used for measuring. Seventy-four percent of the eighth graders (as compared to 41 percent of fourth graders) knew that a scientist would have the most trouble measuring how much better one dog is than another; the other answer options related to how tall or how heavy an object is, how strong a person is, and whether a person has a fever. Over half of the grade 4 students knew that a pencil shown in the item was approximately 10 centimeters long as opposed to other distractors presented in centimeters. Seventy percent of the eighth graders recognized that the pencil was approximately ten centimeters long as opposed to ten of some other metric unit.

Only 48 percent of the eighth graders and 60 percent of the eleventh graders knew that when scientists carefully measure a quantity many times, they expect that "most of the measurements will be close but not exactly the same." Twenty-four percent and 20 percent of the students at those grade levels believed that scientists expect all of the measurements to be exactly the same. (See sample test item number 3 in the appendix.) Those responses are consistent with many students' misconceptions about science dealing only with facts, exactness, etc. On another eleventh grade item, three-fourths of the students knew that the average of 30 measurements of an object would be closer to the actual result than the smallest or largest result, or the average of the ten smallest or ten largest results.

Exhibit 3 below shows the results on two items on which the grade 11 performance leaves a great deal to be desired. Two things are involved in the first--rounding to the nearest centimeter and recognizing that the left end of the pencil is aligned with "1" not "0" (the desired procedure for using a ruler). The grade 4 result on the second item is not surprising, although by grade 4 most students should have been exposed to multiunit gradations.

Exhibit 3

Applying Measurement Skills

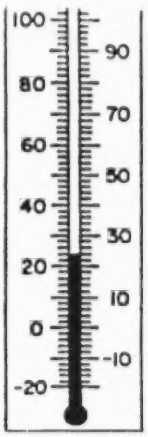


Percent Correct Grade 11		
19	*	A. 9 cm
25		B. 9.2 cm
17		C. 10 cm
36		D. 10.2 cm
1		E. 92 cm
1		F. 102 cm

What is the length of the pencil to the nearest centimeter?

Percent Correct Gr.4	Gr.11		
3	3		A. 17^{o}
74	35		B. 22^{o}
21	61	*	C. 24^{o}
2	1		D. 27^{o}

What temperature is shown on this thermometer?



Interpreting/Translating Data

Most of the test items in this category involve reading or translating information presented in graphic or tabular form. Generally, the performance of grade 4 and grade 8 students in this area was strong. Fifty to ninety percent of the fourth graders responded correctly to items involving bar graphs, line graphs, and tables. Of course, performance was higher on direct reading tasks than on items requiring some interpretation. The results

depicted in Exhibit 4 are quite good considering the item requires more than simple reading of data.

Exhibit 4

Interpretation of a Table--Grade 4

A doctor kept records of breathing rates of people when they were resting. He made the chart below.

BREATHING RATES

Person	Breaths in a Minute
Baby boys	38
7-year-old girls	25
7-year-old boys	25
10-year-old boys	20
Mothers	16

Percent Correct Gr. 4	The chart suggests that
8	A. boys breathe faster than girls.
4	B. girls breathe faster than boys.
8	C. older people breathe faster than younger people.
72	*D. younger people breathe faster than older people.
7	E. I don't know.

Eighth graders' performance in this area was also relatively strong. Two-thirds of the grade 8 students could correctly associate tabular data with a corresponding line graph. Another two-thirds understood the meaning of two lines intersecting in a more complex graph.

Although the test items administered to grade 11 students were considerably more complex, performance was lower than expected. One item described how many eggs a particular type of creature produced at a time and how many times a year it does so. It asked the students how many creatures there would be in a year, assuming there were no deaths. The students' responses to this question illustrated that most students have little concept of exponential growth. Performance on a follow-up item asking the students to pick the graphic representation of this growth was consistent with this conclusion. Only 24 percent of the eleventh graders selected the appropriate graph, while almost half selected a graph showing a linear increase.

Just under three-fifths of the students successfully interpreted a table in which two variables showed an inverse relationship to one another. Just over three-fifths successfully read a wind chill table--many students erred by ignoring the signs for above zero and subzero temperatures. On an item involving smoking and lung cancer, only 44 percent of the eleventh graders recognized that the strongest evidence of a link between the two would involve direct observation as opposed to other forms of evidence.

Drawing Conclusions/Inferences

Performance in this category was somewhat mixed. Grade 4 students did well (69 percent) in recognizing that if a person dropped a piece of wood in water and it sank, the best that could be concluded is that "some pieces of wood sink in water." On a multistep exercise administered at grade 8, students had to refer to a line graph and to a table to determine when a dog's daily allowance of food should have been increased because of the weight of the dog. Three-fifths of the students succeeded at this.

On a grade 8 and 11 item, a graph illustrated the relationship between crop yields and rainfall. The graph clearly showed a curvilinear relationship in which crop yield increased as rainfall increased only to a certain point, and then declined as rainfall continued to increase. Fifty-five percent of the eighth graders and 68 percent of the eleventh graders successfully interpreted the graph. The most popular distractor at both grade levels simply stated that the more rainfall, the better the crop would grow.

A pair of items based on a table relating automobile stopping distances to speed required eighth and eleventh grade students to interpolate and extrapolate. (See sample test item number 4 in the appendix.) Fifty-one percent of the eighth graders and 67 percent of the eleventh graders succeeded on the former, while 34 percent of the eighth graders and 46 percent of the eleventh graders succeeded on the latter task.

On an open-ended or free-response question, 44 percent of the eighth graders and 63 percent of the eleventh graders successfully attributed faster weight gain in hamsters to a particular diet based on tabular information. Of greater concern than those percentages on this item are the non-response rates -- 19 percent at each grade level. On a similar item at grade 4, the non-response rate was only 12 percent. However, those students were not nearly as successful. The data was presented in the form of a bar graph relating heart beats per minute to type of exercise. Only 13 percent of the fourth graders wrote a valid conclusion (e.g., the heart beats faster when running). Seventeen percent gave an unwarranted conclusion (e.g., walking is best), and another 16 percent simply described what the graph is about, but did not draw a conclusion. While open-ended exercises are generally more difficult, students clearly need more practice at drawing their own conclusions rather than selecting conclusions from a set of alternatives.

Life Sciences

Characteristics of Life

Grade 4 students show only a fair understanding of the differences between living and non-living things. While 68 percent of the students recognize that a wooden chair is made from material that was once alive, the remaining 32 percent indicated that a metal cabinet, a stone wall, or a glass window was once alive. Although basic life functions are frequently covered by grade 4, the performance on one assessment item would suggest that students don't grasp their meaning. Eighty-one percent of the students mistakenly indicated that not all living things can take in materials from the surroundings, eliminate waste materials, reproduce, or grow. Fortunately, by grade 8, the situation

changed. On a slightly more difficult item using biological terms for the functions, 73 percent of the eighth graders successfully identified locomotion as not being a basic life function of all living things.

Just over half of the eighth graders knew that the cell is the simplest biological unit, as opposed to an organ, tissue, organism, or population. Eighty-one percent of the eleventh graders succeeded on this item. On a more difficult item, over two-thirds of the eleventh graders recognized that a group of similar cells doing the same thing is called a tissue instead of an organism, an organ, or a system.

Heredity/genetics appears to be an area of weakness for Connecticut students. Only 47 percent of the eighth graders knew that if the parent animals are larger than most other animals of the same species, their offspring will probably be larger also. Only half of the eleventh graders knew that a new cell has the same number of chromosomes as the parent cell. Fewer than one-third could determine the fractions of offspring with different traits given characteristics of the parents and information on the dominant trait.

Stability of populations is a concept not understood by many eighth or eleventh graders. Only 25 percent and 41 percent of the students at those grade levels respectively recognized that the number of weeds in a field one year would be similar to the number in the field in previous years. The question indicated how many seeds were produced in one year, and many students felt that that figure was an indication of the number of weeds that would result.

Animal Life

Although most of the test items in this category dealt with human biology, a few dealt with non-human forms. Most fourth graders (93 percent) knew that a caterpillar (shown in a picture) would grow up to look like a butterfly (also shown in a picture). Only two-fifths of these students, however, knew where a baby chick growing inside an egg gets its food. At grade 8, 57 percent of the students knew that humans have probably been on earth the shortest time, as opposed to alligators, dragonflies, fish, or snails. An open-ended question asked these students to list three ways in which some animals have adapted to meet the challenge of changing seasons. Sixty-four percent listed at least two correct responses. Twenty-nine percent of the students' lists included migration, 57 percent included hibernation, 43 percent included winter coats, and 29 percent included the storing of food.

Items covering human biology addressed such topics as nutrition, parts of the body, disease, and reproduction. Only 47 percent of the fourth graders knew that foods pictured in the item were important because they contain an essential vitamin. Almost an equal number indicated that the foods pictured were protein foods, used for building muscle. Fifty-three percent of the fourth graders could identify the most balanced menu from a set of four menus. At grade 8, 36 percent of the students knew that lean meat provided the most protein, as compared to boiled potatoes, green beans, oatmeal, and white bread. Twenty percent of these students selected green beans, the second best answer. Clearly students have some exposure to food groups, but they have some difficulty linking particular foods to their groups.

Table 5 lists facts tested by selected human biology items and provides the percent correct for each. The reader is cautioned that the way an item is actually worded can have a major influence on performance. The technical report for this science assessment includes a listing of all of the assessment items as they actually appeared in the booklets.

Table 5

Results on Selected Human Biology Items

Fact Tested	Percent Correct Grade 4	Grade 8	Grade 11
The most important gas taken into lungs is oxygen.	81	-	-
The most important gas taken into blood stream during breathing is oxygen.	-	83	-
The main function of the heart is to circulate blood.	70	-	-
An important function of the kidneys is to remove waste from blood.	-	63	-
The main function of kidneys is to regulate composition of body fluid.	-	-	59
White blood cells fight disease.	-	70	78
Veins carry blood <u>to</u> the heart <u>from</u> other parts of the body.	-	-	31

Students seem to have a good general knowledge of the parts of the body; however, the results on the grade 11 items about kidneys and veins suggest that students do not have a deeper understanding of how the human body works.

Test items covering emergency first aid procedures suggest that there are some myths that need dispelling. Only 43 percent of the fourth grade students recognize that covering a wound with a cloth and pressing firmly is the appropriate procedure for dealing with a rapidly bleeding cut. Forty-eight percent of the students felt that the first thing that should be done is to wash the cut carefully. The emphasis on the importance of cleanliness would result in inappropriate behavior in this instance. Sixty-four percent of the eleventh graders identified the appropriate behavior for dealing with a person who is choking on food. Twenty-eight percent incorrectly responded that the appropriate action would be to quickly pound the person on the back.

Eighty-one percent of the eighth graders knew that a person with diabetes cannot properly use sugar; however, only 26 percent knew that cancer is uncontrolled cell division. It is disappointing that only 46 percent of the eleventh grade students knew that antibiotics are capable of controlling bacteria. Twenty-eight percent believed that antibiotics control viruses.

Only 56 percent of the students in grade 8 and grade 11 recognize that bacteria benefit living things in all of the ways listed in the question--i.e., causing decay, making serums, and producing certain foods.

Human reproduction is another area in which students seem to have a general knowledge, but lack specifics. Three-quarters of the eighth graders know what fertilization is; however, only 27 percent of the eleventh graders know that the egg is released 14 days after menstruation. Only 48 percent know the function of the placenta.

Plant Life

At the fourth grade level, 67 percent of the students identified the picture of the cocklebur as the seed that would be spread by sticking to the fur of animals, and 76 percent knew that pine trees stay green during the winter. At all three grade levels, performance was poor on items dealing with parts of the plant and photosynthesis. At grade 4, while 89 percent of the students identified the roots as the part of the plant holding the plant in the soil, only 44 percent knew that seeds were produced in the flowers, only 40 percent knew that the leaves take in sunlight and produce food, and only 39 percent knew what is inside a seed. At grade 8, only 57 percent of the students knew that the seeds come from the flower of the plant. Only 45 percent of the eighth graders and 57 percent of the eleventh graders knew that cell walls distinguish plants from other organisms.

While 53 percent of the fourth graders knew that plants provide food and oxygen to animals, only 60 percent of the eleventh graders demonstrated that same knowledge on an item parallel to the fourth graders' item. When asked to identify the most important result that the death of all green plants would have on humans, only 37 percent of the eighth graders indicated that sooner or later humans would die of starvation. Thirty-six percent simply indicated that humans would get a little sick because we could not get certain vitamins.

Only 24 percent of the eighth graders and 32 percent of the eleventh graders identified sugar and oxygen as two products made by plants during photosynthesis. (See sample test item number 6 in the appendix.) An open-ended item asked the students to name three things plants need to produce food. Seventeen percent of the fourth graders and 28 percent of the eighth graders succeeded. Two-thirds to four-fifths of the students' lists included water and light, and one-quarter to two-fifths included air, but very few included chlorophyll.

Ecology/Environment

Performance in this category indicates that students have only a superficial knowledge of ecological and environmental issues. Some items did show acceptable performance. For example, 47 percent of the fourth graders recognized that depletion of nutrients in the soil was a problem when plants did not grow as well year after year in a particular garden. Fifty-seven percent of the eighth graders and three-quarters of the eleventh graders knew that fish and ocean animals use oxygen produced by plants living in the water.

On an open-ended exercise, only 19 percent of the fourth graders and 50 percent of the eighth graders were able to provide an example of a food chain

comprised of at least three links. An eleventh grade item described a meadow in which there were many rabbits that ate grass, and there were many hawks that ate the rabbits. Two-thirds of the eleventh graders knew that if a disease killed a great number of the rabbits, the grass would grow taller and the hawk population would decrease. This item could actually be viewed as a simple logic exercise not requiring instruction; many more students should have succeeded on it even without specific instruction in the content. (See sample test item number 7 in the appendix.)

An open-ended question asked the eleventh graders to discuss several things that might occur if some strange fish from another part of the world were released in a Connecticut lake. Only 39 percent of the students named two or more correct and distinct effects. Over three-quarters of the lists included the idea that the strange fish might not survive. However, only 25 percent included "the killing of other fish," and only 9 percent included "competing for food." One-quarter of the students made a general reference to the disrupting of the ecological balance. Thus, the majority of the students came up with the simple response about the fish not surviving, but very few recognized the bigger issue--the effect on the larger system. On a multiple choice item, only 50 percent of the eleventh graders recognized that human beings thin forests, spray insecticides and herbicides, remove undergrowth, and plant new trees to make forests produce more of what humans want. The other 50 percent of the responses were well distributed among other reasons--to protect humans and animals from disease, to make a better home for wildlife, and to protect forests from forest fires. These results might suggest a shared attitude among students that humans and science can do a better job than mother nature.

A multiple-choice item asked the students to indicate which of the following would cause a river to become polluted--fertilizer from farms, soapy laundry water, and chemical wastes from factories. Thirty-five percent of the fourth graders, 61 percent of the eighth graders, and 71 percent of the eleventh graders correctly answered "all of the above." The most popular distractor was "chemical wastes from factories," many students apparently seeing only the beneficial characteristics of fertilizer and soap. Table 6 shows results on items about particular types of pollution. Notice that only 28 percent of the eighth graders identified cars as the major cause of air pollution in most large American cities. Sixty-three percent identified factories--the producers of more visible pollutants. One would expect stronger performance on the items about acid rain and the greenhouse effect because they are relatively current topics for instruction and for the media. However, it may be that the media has a negative influence on students' understanding of the problems associated with nuclear power. The low percentages of students at grades 8 and 11 identifying elevated water temperature as the most likely harmful effect of nuclear reactors relates inversely to the large numbers of students choosing radioactivity. (See sample test item number 8 in the appendix.) The harmful effects of radioactivity are emphasized by media reports on nuclear power. In a question asking students to name the best reason for restrictions on the sale of tuna containing high concentrations of mercury, over two-thirds of the eleventh graders knew that mercury accumulates in humans and that over time it could build up to lethal levels.

Table 6

Results on Selected Multiple-Choice Items about Pollution

Fact Tested	Percent Correct Grade 8	Grade 11
Major cause of air pollution in cities is cars.	28	-
Sulfur dioxide from factories falls as acid rain.	54	-
The most likely harm caused by nuclear reactors is the high elevated temperature of water.	49	57
The primary direct outcome of the greenhouse effect is higher temperatures.	-	66

Earth and Space Sciences

Astronomy

The assessment results suggest that knowledge of specific facts is covered well in instruction of astronomy. Over four-fifths of the students at grade 4 knew that day and night occur because the earth turns. Just under two-thirds of these students knew that the sun was the largest of the bodies listed in one item, and just over two-thirds knew that the sun is a star. Approximately four-fifths of the fourth and eighth graders knew that the earth moves around the sun each year. Thus, students appear to have a fairly good grasp of many basic facts in astronomy. Notable exceptions were revealed by questions about the time it takes the moon to run through its phases (related item answered correctly by 25 percent and 50 percent of the fourth and eighth graders respectively) and to the fact that reflecting light is a common characteristic of planets.

Questions requiring higher thought processes yielded considerably lower results. Asked why telescopes are often built on mountain tops, only one-fifth of the fourth graders and half of the eighth graders responded correctly (i.e., to get above smog, dust, and fog). Three-fifths of the younger students, and one-third of the grade 8 students, indicated that the reason telescopes were so located is so that they would be as close as possible to the stars. These students clearly have no concept of relative distances in space. (See sample test item number 9 in the appendix.) Poor reasoning skills were evidenced by an item included in the grade 4 and 8 assessment. Only one-quarter of the younger students and one-third of the older could determine that if a person looks west and observes a full moon near the horizon, then it must be about sunrise. Larger numbers indicated that it would be about sunset.

When asked which phenomenon is NOT a cause of seasonal changes, only thirty-three percent of the eighth graders and 39 percent of the eleventh graders correctly answered "the rotation of the earth on its axis." This result is surprisingly low since students are commonly taught the information contained in the other distractors--i.e., the tilt of the earth relative to the

sun, the revolving of the earth around the sun, and the angle at which the sun's rays strike the earth. Even more surprising results were obtained on an item about the motion of stars. Only 31 percent of the eighth graders and 35 percent of the eleventh graders correctly indicated that if the stars were watched all night, they would appear to move across the sky as the sun does during the day. Large numbers of students claimed that the stars would stay in exactly the same place!

Climate and Weather

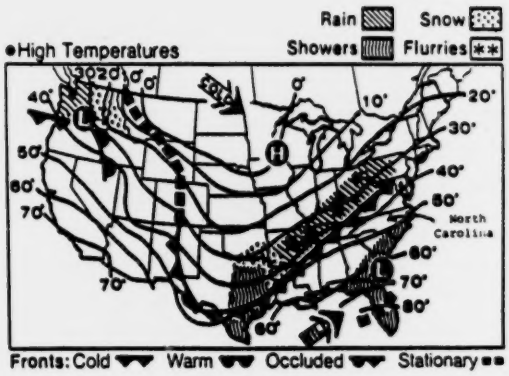
The results on questions about climate and weather are consistent with those discussed in the previous section on astronomy. Students are better in their knowledge of facts than they are with meaning. Over half of the grade 4 students knew that frozen rain is sleet; 83 percent knew that the most damaging common characteristic of hurricanes and tornadoes is high winds. Most fourth and eighth graders (83 and 89 percent respectively) can pick the description of clouds most likely to bring rain. However, only 49 percent of the eighth graders (compared to 78 percent of the eleventh graders) knew that atmospheric pressure is measured by a barometer as opposed to other instruments listed.

As in astronomy, questions addressing the comprehension of concepts in the area of weather and climate produced low percentages of correct responses. Only a third of the eighth and eleventh graders knew why there is less variation between high and low temperatures in coastal areas than in desert areas. Exhibit 5 shows results on items dealing with the meanings of things students are exposed to almost every day. Too few eleventh graders could predict the weather for an area given a weather map showing some very basic principles. Students could arrive at the correct answer in two ways--by recognizing that our weather conditions follow a general west-to-east pattern of movement or by applying knowledge of the fact that air moves from areas of high pressure to areas of low pressure (see Item A). Interestingly, only 42 percent of the eleventh graders succeeded on an item testing that fact directly: the item asked about the direction the wind blows and the answer mentioned high to low pressure areas. Results on the Item B in Exhibit 5 show that most students have no idea what is meant by a weather report indicating that "the relative humidity is 50 percent." A comparable result was obtained on a similar question for eleventh graders dealing with a 20 percent chance of rain.

Exhibit 5

Understanding of Weather Map and Relative Humidity

Item A



Percent Correct Gr. 11		
		Using today's weather map above, predict tomorrow's weather for North Carolina.
16	A.	getting warmer
59	*B	getting colder
18	C.	no change
7	D.	I don't know.

Item B

Gr. 8	Gr. 11		
			The statement that the relative humidity is 50 percent means that
21	15	A.	the chance of rain is 50 percent.
7	8	B.	the atmosphere contains 50 pounds of water per cubic mile.
9	12	C.	the atmosphere contains 50 grams of water per cubic meter.
9	7	D.	the atmosphere would be saturated with water if the air temperature were 50° F.
31	40	*E.	the atmosphere contains half as much water as it could contain at its present temperature.
22	18	F.	I don't know.
1			no response

Geology/Oceanography

There were mixed results on items dealing with the formation of the earth. Over half of the eighth graders knew that the earth's molten core was formed prior to the formation of continents and clouds, and before volcanic eruptions. Fifty percent of the eighth graders and 57 percent of the eleventh graders knew that places where tectonic plates meet are characterized by frequent earthquake activity. These results are generally reasonable, although

there was little growth between grades 8 and 11. Very few eighth graders (27 percent) knew that most caves are formed by the action of underground water on limestone, as opposed to other materials.

Students appear to have some familiarity with fossils. Seventy-eight percent of the fourth graders knew that scientists can learn about animals living long ago by studying fossils. Sixty-eight percent of the eleventh graders, given a list of items that could be found in a rock, knew that raindrop imprints would not be considered fossils. An item requiring the students to link ideas related to the formation of the earth and fossils produced poor results. Only 29 percent of the eighth graders knew that a fossil of an ocean fish found in a mountain outcrop means that the mountain was raised up after the fish died. Thirty-eight percent of the students at that grade level indicated that the fossil fish was carried to the mountain by a great flood. (See sample test item number 11 in the appendix.)

Fifty-one percent of the fourth graders believed that gasoline is the most plentiful fuel in the United States. Only 18 percent of the fourth graders correctly indicated that coal was most plentiful. However, by grade 11 these percentages are almost reversed; 55 percent of the older students recognized that coal is the most plentiful fuel in the United States. Forty-five percent of the eighth graders and 62 percent of the eleventh graders knew that the United States gets most of its energy from oil. Thus, by the eleventh grade, many students have a reasonable knowledge of important facts about natural resources. However, results were not as good on an item asking eleventh graders what is wrong with the idea of switching to electricity as a solution to coal, gas, and oil shortages. Only 38 percent of the eleventh graders selected the correct response--i.e.most electricity is produced from coal, gas, and oil. (See sample test item number 12 in the appendix.) Once again, many students have a lot of facts at their disposal, but do not really understand them well enough to apply them.

Over half of the fourth graders knew that as a diver goes farther below the ocean's surface, the water gets colder and the pressure becomes greater. Students' knowledge of tides and waves is somewhat lacking. Only 35 percent of the eighth graders attributed most waves on the ocean's surface to the wind. Thirty-eight percent of the students attributed most waves to the moon's gravity, and 22 percent to shifts in the ocean floor. On an item dealing more directly with tides, only 53 percent of the eighth graders and 58 percent of the eleventh graders knew that the most direct cause of tides on earth is the effect of the moon's gravitational pull.

Knowledge of the earth's water cycle is generally regarded as important for students. However, only 34 percent of the eighth graders knew that the source of energy for the water cycle is radiation from the sun. Fewer eleventh graders (29 percent) knew that fact. More eighth and eleventh graders identified the earth's rotation as the source of energy for the water cycle than the correct response. Two of the test questions about ocean life produced mixed results. Sixty percent of the eighth graders (compared to 30 percent of the fourth graders) knew that the earliest life on earth probably appeared in the ocean. A grade 8 and 11 item asked students to identify a true statement about salt marshes along the Connecticut coastline. The results were poor, with only 28 percent of the eighth graders and 33 percent of the eleventh graders recognizing that the salt marshes are major breeding areas for many

kinds of ocean life. More students at grade 8 and the same number of students at grade 11 chose as true a statement indicating that the salt marshes would have little life because the water is too salty for fresh water life and not salty enough for ocean life.

Physical Sciences

Physics

At grade 4, 73 percent of the students identified gravity as the force pulling things toward the center of the earth. Sixty-eight percent of the eighth graders knew that as two objects move farther apart in space, the gravitational pull between them gets weaker. Thus, many students have some rudimentary knowledge of forces between bodies. A grade 11 item showed pictures of a wheeled wagon with the handle in various positions, and asked which position of the handle would make the wagon experience the stronger pull to move forward. Many grade 11 students were not familiar with vector components of forces; only 39 percent chose the correct picture in which the handle was extended horizontally. Forty-one percent chose the picture of the wagon in which the handle was angled upward.

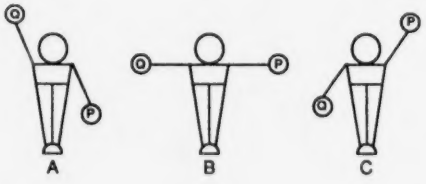
Just over half of the eleventh graders (and 28 percent of the fourth graders) knew that sand is used on icy roads to create friction between car wheels and the road. Over a third of the eleventh graders and half of the fourth graders claimed that the purpose of the sand is to help melt the ice. Clearly, many students are confusing sand with salt.

Items dealing with objects in motion generally produced poor results. Exhibit 6 shows the results on two such exercises. The reader should take special note of results on the first item involving falling objects. Although performance at grade 4 was better than it was at the other two grade levels, the results on the practical tests (see page 32) indicate that these students do not really understand the principles involved. While the second item showed substantial "growth" across grades, one would expect better performance at all levels on an item that students could easily relate to personal experience. On an eleventh grade item (not shown), only 39 percent of the students knew that the speed of a falling object would be greater the farther the object falls. Twenty-nine percent indicated that the acceleration would be greater. This and other items from the assessment clearly suggest that many students had difficulty distinguishing between speed and acceleration.

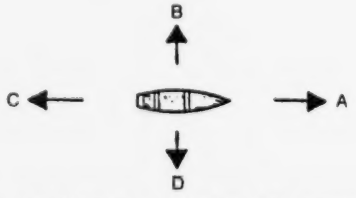
Exhibit 6

Results on Selected Items in Mechanics

Percent Correct Gr.4	Gr.8	Gr.11	
			Suppose that you want to drop a penny and a quarter at exactly the same time and have them hit the floor at exactly the same time. Which picture BEST shows how you would hold the penny and the quarter just before you drop them?
17	31	30	A.
71	56	57	*B.
8	8	9	C.
4	3	3	D. I don't know.
1	1		no response



			A person dives out of a nonmoving boat in the direction indicated by Arrow A. Which arrow shows the way the boat would move?
52	27	13	A. Arrow A
10	7	6	B. Arrow B
31	59	75	*C. Arrow C
6	5	5	D. Arrow D
2	1	1	no response



Heat transfer is not understood by many students. Only 38 percent of the eleventh graders could identify the reason why a spoon handle sticking out of a cup of hot water also gets hot. The correct response was that the molecules bump into each other. The most popular distractor suggested that an electric current was started in the handle of the spoon. Despite their lack of the more formal knowledge of heat transfer, many students still recognize that insulating walls and attic spaces is the most effective energy-saving measure in a home, compared to other measures listed. This item was answered correctly by 40 percent of the fourth graders, 73 percent of the eighth graders, and 69 percent of the eleventh graders.

Test results indicate that most students have been exposed to the notion of light-colored objects reflecting more light than dark-colored objects. Forty-two percent of the fourth graders and 77 percent of the eighth graders

correctly selected the reason why it is cooler to wear light-colored clothes in the summer.

Eighth graders did better than expected on an item about communicating on the moon. Fifty-eight percent of the students correctly indicated that ringing a bell would not be an effective way to communicate; conversely, they realized that flashing a light, using a radio, or waving would present workable alternatives. Only 27 percent of the eleventh graders recognized that higher frequencies of a wave correspond to shorter wavelengths.

Ninety-one percent of the grade 4 students knew that a magnet would pick up an iron nail as opposed to other objects; only 37 percent knew that magnetism, rather than heat, gravity, or air pressure, is what allows us to use a compass. Fifty-nine percent of the fourth graders and 72 percent of the eighth graders correctly identified copper as the best conductor from a list of different materials. It was somewhat disappointing to find that only between half and two-thirds of the students in grades 8 and 11 realized that electricity is a flow of electrons and not other particles.

On an item asking students to identify the picture of a circuit that would result in a bulb lighting when the switch was closed, the percents correct at the successive grades were 46 percent, 65 percent, and 75 percent. While there was consistently higher performance at successive grades, one would expect that something as basic as a simple closed circuit should be understood by more students. Interestingly, practical test results show that over 90 percent of the students at all grades could set up a simple circuit that worked given the materials (bulb, battery pack, wire) and time to try different approaches.

Another fact of practical importance was known by half of the eighth graders and two-thirds of the eleventh graders--namely, that an electric clothes dryer would consume more energy in 15 minutes of continuous operation than a television, a vacuum cleaner, or an electric can opener. (See sample test item number 14 in the appendix.) Just over a third could identify the reason for a fuse being used in an electric circuit. If energy conservation and safety in the home are concerns, students should have a much better knowledge of various practical concepts in electricity.

Chemistry

The results depicted in Table 7 show that knowledge of the nature and structure of matter is an area of particular weakness for all students. On the average, they show very little understanding of atoms, elements, and compounds. Of particular interest are the results on the item dealing with the conservation of matter (also shown in Table 7). While it is reasonable to expect that most students have heard that "matter can neither be created nor destroyed," too few students realized that the situation described in the test item is one related to that law of conservation of matter.

Several assessment questions dealt with states of matter. Sixty-four percent of the grade 4 students knew that water changing from a liquid to a gas is evaporation. Forty-four percent of the fourth graders knew that water drops appearing on the outside of a cold glass of lemonade came from water in the air. However, on an open-ended question administered at grade 11, only 38 percent of the students could correctly explain how the outside of a glass of

iced tea sometimes gets moist. Responses were scored as correct if they mentioned the word condensation as well as if they simply explained that water vapor in the air turned to a liquid. On another grade 11 multiple-choice item, 56 percent of the students recognized that the various states of water differ in the average speed at which the molecules are moving. Fewer than a third of the grade 11 students knew that an increase in the pressure of a gas could result from a decrease in the volume.

Table 7

Results on Selected Multiple-Choice Items about the Nature/Structure of Matter

Fact/Concept	Percent Correct		
	Grade 4	Grade 8	Grade 11
Objects in list are all matter except "ideas."	43	--	--
Plastic, wood, iron are all made of atoms.	39	--	--
Part of atom with positive charge is proton.	--	47	--
Neutral atom losing electron forms ion.	--	25	--
Neutral oxygen atom contains protons, neutrons, and electrons.	--	--	49
Putting sand and salt together makes a mixture.	50	--	--
Water is a compound.	--	28	50
Pure substance that can't be decomposed by several means listed is most likely an element.	--	31	53
1 kg salt dissolved in 20 kg water makes 21 kg saltwater. (conservation of matter)	44	51	49
Object on earth and moon has same mass.	--	28	34

Several assessment exercises covered chemical changes. Forty-five percent of the fourth graders could identify a change (a candle burning) in which a different substance is formed. Sixty-nine percent knew that iron would be more likely to rust when it is damp. Sixty-eight percent of the younger students knew that the burning of a campfire could be speeded up by blowing on the fire. Seventy-one percent of the eighth graders could explain why fanning makes a fire burn better--i.e., it increases the supply of oxygen.

As one would expect, nonresponse rates were fairly high (21 to 35 percent) on items requiring formal instruction in high school chemistry. Only 37 percent of the eleventh graders knew that chemical changes could be explained in terms of the gain, loss, or sharing of electrons. Only 39 percent knew

about catalysts. Fifty-nine percent of the students could supply the coefficient that would correctly balance a chemical equation.

CHAPTER IV

RESULTS OF THE PRACTICAL TESTS

As explained in the first chapter, several practical exercises were administered to a randomly selected subset of the statewide sample of students. The exercises required students to measure and observe things, manipulate objects, and conduct simple investigations. Thus, the practical tests measured not only what students already knew, but also what they could do or figure out. Three hundred students were involved in the practical testing at each of the three grade levels. They were tested in pairs by trained administrators, the two members of each pair completing different sets of exercises.

Grade 4 Results

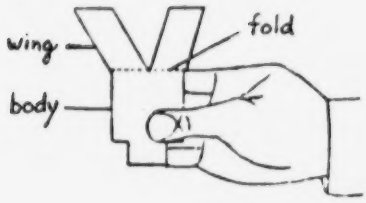
Since sorting and classifying is an early primary grade activity, it is not surprising that the fourth graders did well at sorting objects. The objects involved were nuts, bolts, and washers; they were either large or small, and made of either brass or steel. Ninety-three percent of the fourth graders correctly sorted the objects into two mutually exclusive groups. Eighty-six percent correctly sorted a second time into two different groups. Ninety-two percent correctly sorted the objects into three mutually exclusive groups.

The fourth graders also did well on some simple measurement tasks. Asked to measure the length of a wooden block to the nearest centimeter, 91 percent of the students were successful. The other 9 percent of the students gave responses to one decimal place. Asked to read the temperature of a container of cold water and a container of hot water, 85 percent of the students gave measurements within two degrees of the administrators' readings. The thermometers were marked in one-degree gradations. Another question required the students to determine the difference between the two water temperatures. Eighty-six percent of the students correctly computed differences based on their readings. A follow-up question asked the students to guess what the temperature would be if they mixed a larger quantity of cold water with a lesser quantity of warm water. One-half of the students correctly guessed a temperature between the warm and cold water temperatures and closer to the cold water temperature. Another 24 percent guessed temperatures in between the two extremes, but not closer to the cold water temperature.

The students were given a battery pack with wire leads, a light bulb with wire leads, and insulated wires with alligator clips for connectors. They were instructed to use these materials to make the bulb light. Eighty-five percent of the fourth grade students succeeded. They had ample time to try out different configurations. Interestingly, on a multiple-choice item requiring the students to identify the picture of a simple circuit in which the bulb would light, only 46 percent of the fourth graders succeeded. Since most students can obviously figure out how to hook up the simple circuit, one might expect that giving the students more practical experience of this sort with opportunities to observe and discuss the scientific principles involved would lead to a higher success rate on related test items.

The students were also given a pair of paper "helicopters." (See Exhibit 7.) The students were asked to determine if there was anything different about the way the two helicopters fell. Seventy-eight percent correctly noted that they fell at different speeds or with different rates of spin. As a follow-up question, the administrators asked the students why they fell differently; 73 percent of the students correctly attributed the different motions to wing length--the only difference between the two helicopters.

Exhibit 7. Paper Helicopter



Another simple investigation provided the students with a timer and an apparatus enabling them to set up pendulums with bobs of different weights and varying string lengths. Asked to time 40 complete swings of various pendulums, 89 percent of the students timed them reasonably accurately--i.e., they started and stopped the timer and counted swings appropriately. After being directed through the timing of the various pendulums, the students were asked what they could conclude about pendulums. Sixty-five percent correctly concluded that shorter pendulums swing faster. Eight percent concluded the opposite, and 24 percent gave conclusions not comparing speeds.

When asked what would fall faster when dropped, a penny or a quarter, only 5 percent of the fourth graders indicated that both would fall at the same rate. This is quite inconsistent with the results on a multiple-choice question in which 71 percent of the fourth graders correctly identified the picture showing how the penny and quarter should be held so that they would hit the floor at the same time when released. On that item, many more fourth graders answered correctly than eighth and eleventh graders. There might be any number of interpretations of this inconsistency. One that might be considered is that on the paper-and-pencil test, the fourth graders responded with what they had been taught, not what they really believe. When presented with the question prior to actually testing out their answer on the practical test, one suspects they may have been expressing their true belief--in other words, they really did not believe what they have been taught. Another possible explanation of the paper-and-pencil results is that the fourth graders read less into the question and chose the figure with the penny and quarter held at the same height because it seems right if you don't realize that pennies and quarters weigh different amounts. The older students, however, consider the weight factor, which might lead them astray. In conducting the investigation to determine what would actually happen, 94 percent of the students correctly controlled the height of the two coins when released.

Grade 8 and 11 Results

Asked to measure the length of a block of wood to the nearest centimeter, 88 percent of the eighth and eleventh graders gave the correct response. However, an additional 4 percent at grade 8, and 7 percent at grade 11, gave unrounded responses accurate within one-tenth of a centimeter. Thirty-five percent of the eighth graders and 34 percent of the eleventh graders correctly reported the mass of the block to the nearest gram. An additional 17 percent of the eighth graders and 19 percent of the eleventh graders reported unrounded

answers accurate within one gram. Approximately 70 percent of the eighth graders and 40 percent of the eleventh graders reported having used a triple-beam balance five or fewer times. Twenty-nine percent of the eighth graders and 17 percent of the eleventh graders reported never having used a triple-beam balance (see Tables 8 and 9). Thus, the relatively poor performance on the weighing task might well be attributed in part to limited experience in the use of the equipment. Very few students (3 percent at grade 8 and 12 percent at grade 11) were able to determine the density of the wooden block. Given the performance level on the weighing task and knowing the difficulty students have with the concept of volume, it is not surprising that so few could determine the density of the block. If computing volume is valued as a skill, students need more instruction in it.

Using a 100-milliliter graduated cylinder, 45 percent of the eighth graders and 76 percent of the eleventh graders accurately measured 66 milliliters of water. Twenty-nine percent of the eighth graders and 9 percent of the eleventh graders were scored incorrect because they placed the top of the meniscus at 66 milliliters. On the same electric light activity used at grade 4, 93 percent of the eighth graders and 91 percent of the eleventh graders succeeded at getting a bulb to light. Again, these percentages are considerably higher than the percents correct on the multiple-choice question requiring students to identify the working circuit, indicating again that students are willing to work with apparatus when given the opportunity. An additional task for grades 8 and 11 students was the reading of an ammeter--an unusual scale--with the smallest gradations corresponding to 2.5 milliamps. Sixty-three percent of the eighth graders and 40 percent of the eleventh graders gave a reading accurate within one milliamp.

Given a microscope provided by the school and an onion-skin slide, 83 percent of the students at grade 8 and 88 percent of the students at grade 11 were able to bring the slide into clear focus or at least to the point where the focus could be improved only a little. Asked to draw what they saw, relatively small percentages (13 percent at grade 8 and 11 percent at grade 11) produced unrecognizable drawings. The drawings of 28 percent of the eighth graders and 42 percent of the eleventh graders were considered clear, accurate drawings; drawings of the remaining students were spread over various categories defined by varying degrees of clarity and completeness.

Table 8 shows the percentages of responses to questions about the experience of students with various pieces of equipment. Of the apparati listed, students had least experience with the triple-beam balance and the electrical circuit.

Table 8

Students' Experience with Laboratory Equipment

Experience with:	Triple-Beam Balance		Graduated Cylinder		Setting up Electrical Circuit		Microscope	
Percent at:	Gr.8	Gr.11	Gr.8	Gr.11	Gr.8	Gr.11	Gr.8	Gr.11
Never	29	17	11	4	29	29	3	6
1 or 2 times	23	12	25	9	25	21	13	3
3 to 5 times	17	10	22	12	19	12	18	5
6 to 10 times	12	10	13	12	9	13	24	13
More than 10 times	18	49	28	62	16	22	41	69

Table 9 shows the relationship between previous experience with the triple-beam balance and performance on the weighing task. Students with more prior experience were clearly more likely to succeed on the weighing task.

Table 9

Previous Experience and Performance with Triple-Beam Balance

		Grade 8 Experience with Triple-Beam Balance	
		0-2 times	3 or more times
Measured Weight of Block	acceptable	29	48
	not acceptable	49	23

Half of the eighth and eleventh graders involved in the practical testing spent their time conducting an investigation called the "Survival Exercise." Exhibit 8 shows the problem and the materials available to the students.

Exhibit 8

The Survival Exercise

Materials: paper and pencil for recording results and conclusions
2 large containers, one for hot and one for cold water
pins, rubber bands, tape, and clips for fastening things
a timing device
5 cans (4 metal and 1 plastic; 2 metal the same size)
lids for the cans, with openings for thermometers
2 thermometers (Celsius) in lids
1 metric ruler
2 fabrics to be tested (plastic and blanket)
paper towels for any spills

Statement of Problem: Imagine you are stranded on a mountain side in very cold, dry weather. You can choose a jacket made from one of the two fabrics in front of you. This is what you have to find out:

Which fabric would keep you warmer?

You can use any of the things in front of you. Choose whatever you need to answer the question. You can:

- use a can instead of a person
- put warm water inside to make it more life-like
- make it a "jacket" from the material.

Make a clear record of your results and conclusions so that someone else can understand what you have found out.

Many different aspects of experimental design and execution were scored in the Survival Exercise. Design characteristics included the approach used, as well as whether various factors were controlled. Execution was evaluated in terms of measurements made, recording of data, and consistency of conclusion with data.

Table 10 shows the percentages of grade 8 and 11 students using various approaches to solving the problem. If one considers approaches 6 through 9 the more "desirable" ones, then 69 percent of the eighth graders and 70 percent of the eleventh graders opted to use a desirable strategy.

Table 10

Approaches Used to Solve Survival Problem

Percent of Students			
Grade 8	Grade 11		
3	0	1.	fabric wrapped around hand--qualitative
1	1	2.	fabric around thermometer--qualitative
3	2	3.	fabric around thermometer--set temperature drop: time interval measured
1	1	4.	fabric around thermometer--set time interval: temperature drop measured
9	13	5.	fabric around can and water--qualitative
11	16	6.	fabric around can and water--set temperature drop: time interval measured
41	36	7.	fabric around can and water--set time interval: temperature drop measured
3	8	8.	fabric around can and water--same final temperature: time interval measured
14	10	9.	final temperature measured--not timed
13	9	10.	other

Various factors had to be controlled depending on the approach used. These five factors included the cans (size/material), fabric (size and fastening),

water (initial temperature), water (volume), and measurement intervals or temperature drop. Each of these factors was controlled by between half and three-quarters of the students at the two grade levels. The percentages of eleventh grade students who controlled a factor was 8 percent higher than the eighth grade figure on the average. Table 11 shows the correlation between performance on the control aspects of the Survival Exercise and performance on a multiple-choice item dealing with the control of variables in an experiment. Students who answered the multiple-choice question correctly tended to control more of the variables in the performance task.

Table 11

Control of Variables--Performance Test vs. Multiple-Choice Test

		Grade 11 Multiple-Choice Item on Control of Variable (How improve experiment?)	
		Correct	Incorrect
Number of Variables Controlled (Survival Experiment)	4 or 5	29	17
	fewer than 4	14	25

The grade 11 students also performed better than grade 8 students on the "execution" measures. Ninety percent of the eleventh graders, compared to 69 percent of the eighth graders, reported temperature measurements within two degrees of the test administrators' readings. Sixty-five percent of the eleventh graders, compared to 52 percent of the eighth graders, used a reasonable measurement schedule for the detection of a temperature change. Sixty-five percent of the eleventh graders, compared to 58 percent of the eighth graders, organized and recorded data clearly enough to permit appropriate interpretation. Lastly, 57 percent of the eleventh graders, compared to 51 percent of the eighth graders, drew conclusions consistent with their data. (Because of design or execution problems, a third of the eighth graders and a quarter of the eleventh graders found it impossible to reach a conclusion.) The test administrators also provided an overall evaluation of each student's experiment. The design and execution of the experiment for a quarter of the eighth graders and over a third of the eleventh graders were such that one could "trust" the conclusions. The experiments of another third of the students at both grade levels had minor problems which could create some doubt about their conclusions.

This exercise reflects the importance of exposing students to situations requiring them to design as well as conduct experiments. Controlling variables is such an essential agent of conducting experiments that every opportunity should be used to have students design and critique their own experiments.

CHAPTER V

QUESTIONNAIRE RESULTS

Information on student achievement alone does not provide a complete picture of the status of curricular programs. Therefore, the 1984-85 CAEP made use of comprehensive student, teacher, and principal questionnaires. These instruments generated a great deal of data on such factors as student backgrounds and attitudes, teacher experience, classroom practices, and various program and school variables. Selected highlights of these results are reported in this chapter.

Student Questionnaire

At all three grade levels, the students' self-reported grades in science correlated strongly with performance on the CAEP science tests. Students who reported receiving mostly As in science scored from 21 to 28 percentage points higher on the test than students who reported receiving mostly Ds or Fs. While this result should not be at all surprising, it can be regarded as partial evidence of the curricular validity of the test as well as the honesty with which the students answered the questionnaire items.

Apparently, students' attitudes toward science deteriorate as students progress through the grades. Seventy percent of the fourth graders, 57 percent of the eighth graders, and 40 percent of the eleventh graders reported that science was their favorite or one of their favorite subjects. Undoubtedly, science is more fun at the earlier grades. However, it may also be that in the elementary grades, science is taught so seldom compared to other subjects that it is regarded as a special treat. Interestingly, approximately three-fourths of the fourth graders believe that science will be very useful to them when they grow up. Sixty-nine percent of the eighth graders and 57 percent of the eleventh graders either strongly agreed or agreed with a statement about science being useful to them in future work. Fifty-eight percent of the eighth graders and 59 percent of the eleventh graders "strongly disagreed" or "disagreed" with the statement about science not being useful in their future day-to-day lives. Thus, there are still reasonably large percentages of students at the upper grade levels who are not convinced of the importance of science to them.

Students also pursue outside interests in science less in the upper grades. The percentages of students who read science books for pleasure at least a few times per month are 59 percent, 39 percent, and 28 percent at grades 4, 8, and 11 respectively. The percentages of students who watch television shows about science at least a few times a month are 67 percent, 37 percent, and 31 percent across the same grade levels. Much like other variables to be discussed later, reading and watching programs about science exhibit a curvilinear relationship with student achievement. That is, students who pursue these activities a few times per month scored as much as ten points higher than students who reported pursuing these activities never or almost every day.

Eighty-three percent of the fourth graders reported that it is usually true that their teacher cares about how well they do in science. Eighty-seven

percent of the eighth graders and 75 percent of the eleventh graders either "strongly agreed" or "agreed" with a similar statement. More positive views about teachers caring about their work are associated with slightly higher scores on the science tests.

Forty-one percent of the fourth graders, 38 percent of the eighth graders, and 65 percent of the eleventh graders reported that their parents talk to them about their science school work either never or just a few times a year. At grades 4 and 8, such communication shows a curvilinear association with student achievement. Children whose parents speak with them a few times a month about their science did better on the tests than students whose parents rarely, if ever, speak with them or children whose parents speak with them frequently about their science schoolwork. The latter group, of course, would include students who may be having problems in science. Thus, the curvilinear relationship is not surprising. At grade 11, there was less of an association between parental communication and student achievement; however, one must remember that a large number of grade 11 students (42 percent) were not taking science at the time of the assessment, and would therefore have less reason to speak with their parents about science. Throughout the questionnaire, the students in the eleventh grade who were not taking science were reminded to answer in terms of the last science course they took. This may not always have been kept in mind.

Several questionnaire items asked students about the frequency of various classroom activities in science. Table 12 shows the students' responses pertaining to three of those activities as well as the average test scores of the groups of students choosing the different options. Of particular note in the table are the frequent use of the textbook and the infrequent use of small-group laboratory activities at grade 4. With respect to other activities not shown in Table 12, curvilinear associations with test scores were again common. Apparently, the use of a variety of instructional techniques in moderation can be beneficial. Low use of such techniques, or overdependence on them, may be less beneficial. The questionnaire results show that it is rare for students to take field trips or for guest speakers to visit the science classroom; models and games are used very seldom while there is moderate usage of audio-visual equipment -- films, filmstrips, etc.

Table 12

Frequencies of Selected Classroom Activities and Their Relationships to Achievement

Activity	Grade 4 % of stud.	Grade 4 avg. score		Grade 8 % of stud.	Grade 8 avg. score	Grade 11 % of stud.	Grade 11 avg. score
teacher demonstration							
			never	12	49	8	46
never/almost never	24	55	few times/year	12	53	16	52
few times/month	32	55	few times/month	34	53	36	53
1 or 2/week	22	53	few times/week	29	50	27	50
almost every day	22	49	a little most every day	8	50	8	48
			most of class every day	6	47	5	44
read (work from) textbook							
			never	4	49	9	55
never/almost never	14	52	few times/year	5	51	10	58
few times/ month	21	55	few times/month	12	54	19	55
1 or 2/week	25	55	few times/week	28	54	30	49
almost every day	40	52	a little most every day	23	51	17	47
			most class every day	28	48	15	44
small-group labs							
			never	19	46	8	42
never/almost never	59	56	few times/year	13	52	11	47
few times/month	21	53	few times/month	29	51	39	53
1 or 2/week	12	48	few times/week	21	52	32	53
almost every day	9	46	a little most every day	5	52	4	44
			most class every day	13	55	5	48

At grade 8, there appears to be little association between the amount of time spent on science homework and science test scores. Of course, at grade 8, all students are probably taking a science course, and those who spend a great deal of time on science homework could very well be very weak students in science as well as strong students, thus masking an association between time spent on homework and achievement. At grade 11, however, the students putting the most time into science homework would be the stronger students taking science by choice. The questionnaire results show that the grade 11 students who reported spending more than an hour and a half on a typical homework assignment scored approximately 7 percentage points higher than those who are usually not assigned homework and 4 percentage points higher than students who spend less than one-half hour on a typical science homework assignment. Eighty-two percent of the students at grade 8 and grade 11 reported that they are indeed assigned science homework. Over 40 percent of the students at both grade levels report that their typical homework required between one-half hour and one hour of their time.

Table 13 shows the actual/intended science course backgrounds of grade 11 males and females. There are no surprises here, with most students having taken, taking, or intending to take a course in biology and fewer intending to

take courses in chemistry and physics. The course-taking differences between males and females in the physical and earth sciences may help to explain, in part, male/female differences in test performance described on page 8.

Table 13

Science Coursework of Grade 11 Students

Course	Percent of Students Having Taken, Taking, or Intending to Take Course			Course	Percent of Students Having Taken, Taking, or Intending to Take Course		
	Males	Females	All		Males	Females	All
General Science	69	61	65	2nd Year Biology	16	20	18
Earth Science	64	53	59	2nd Year Chemistry	9	5	7
Biology	83	94	89	2nd Year Physics	12	4	8
Chemistry	66	58	62	Other	29	33	31
Physics	54	30	42				

Teacher Questionnaire

In schools participating in the CAEP science assessment, questionnaires were completed by teachers of fourth grade science, teachers of seventh and eighth grade science, and teachers of science in grades 9-12. Throughout this section these groups of teachers will be referred to as grade 4 teachers, junior high science teachers, and high school science teachers.

Twenty-four percent of the grade 4 science teachers, 62 percent of the junior high science teachers, and 69 percent of the high school science teachers are male. Between 80 and 87 percent of the science teachers across the grade levels have nine or more years of teaching experience. Twenty-eight percent of the grade 4 science teachers, 38 percent of the junior high science teachers and 42 percent of the high school science teachers have gone beyond the master's degree in their education (at least master's degree plus 30). The emphasis in both the undergraduate and graduate backgrounds shifts from education for the grade 4 teachers to the natural sciences for science teachers at the higher levels. While almost all grade 4 science teachers have had at least one college course in biology, a third to a half of these teachers have completed no college coursework in chemistry and physics. Of course the junior high and senior high science teachers have considerably more science content in their backgrounds. Most of the grade 4 science teachers are not science specialists.

The science teachers at all three levels were asked about the number of times they have participated in various professional development activities in the past seven years. The activities included college courses in science, college courses in the teaching of science, workshops in science, workshops in the teaching of science, and conferences or institutes in science. Over a third of the grade 4 teachers have not participated in workshops in science or in science teaching in the past seven years. The percentages not participating in each of the other activities are 70 percent or higher. The situation is somewhat different at the other levels where the teachers are science

specialists. The number of teachers at all grade levels who have not participated in each of the listed activities is still surprisingly high. The teachers at the upper grade levels prefer to participate in training in science content.

The grade 4 teachers were asked how many minutes of science instruction their 4th grade students received per week. Thirty-six percent indicated that their students received 80 minutes or less of science instruction. Seventy-one percent indicated that their students received 120 minutes or less.

Regardless of the grade level, between one-half and two-thirds of the teachers are required to use the same instructional materials as other teachers in their schools who teach the same course. Fewer (just over 25 percent) are required to use the same tests for major examinations. Asked about the availability of equipment, 18 percent of the grade 4 science teachers, 35 percent of the junior high science teachers, and 30 percent of the high school science teachers indicated that they had all they needed without having to share. Over half of the teachers at all three grade levels indicated that they must share with other teachers in the school. Twenty-one percent of the grade 4 science teachers, 8 percent of the junior high science teachers, and 16 percent of the high school science teachers indicate that they cannot acquire a lot of the equipment that they need.

At the junior high and senior high levels, teachers have the greatest say in curricular decisions; at grade 4, the decision-making power is spread more evenly across various groups, including teachers, administrators, and curriculum specialists.

At all levels, the teachers indicated that the most valuable resources for information about new developments in education are journals and other professional publications and other teachers. Next on the grade 4 teachers' list of important resources are local inservice programs, principals, and college courses. Also important at the junior high level are college courses, local inservice programs, and meetings of professional organizations. At the high school level, college courses and meetings of professional organizations are high on the list of useful resources.

The data from teachers on the frequencies of various classroom activities confirm information provided by students. Another set of questions asked of teachers at all three levels pertains to the types of questions they ask aloud in class and on written tests. They were asked to estimate the proportion of questions of different types: questions requiring students to repeat memorized facts, questions requiring students to explain something, and questions requiring students to apply their knowledge to arrive at an answer. The students at grades 8 and 11 were also asked about their teachers' questions. Table 14 shows what the teachers and students reported regarding the types of questions teachers ask. Examination of the table reveals that the general pattern is for greater emphasis on lower level questions, both in class discussion and on written tests. The grade 8 and 11 students estimated greater frequency of explanation and application questions than did their teachers.

The teachers were also given a list of factors which may affect science instruction, and ultimately achievement, in their schools. They were asked how much of a problem each factor presented. The factors viewed as causing serious

problems by the largest numbers of teachers are shown in Table 15. As in other CAEP assessments, the primary concern of teachers appears to be the lack of public support and recognition of teachers. The next major areas of concern for science teachers seem to pertain to equipment and budgets.

Table 14

Types of Questions Teachers Ask

Questionnaire Items	Grade 4	Grade 8			Grade 11		
	Percentage of Teachers	Ave. Test Score	Percentage Stud.	Tchrs.	Ave. Test Score	Percentage Stud.	Tchrs.
OUT LOUD DURING CLASS:							
How many of your questions require the students to give back only facts and other information they have **memorized**...?							
almost all	6	48	14	3	46	23	6
more than half	26	51	21	23	50	26	19
about half	38	51	40	44	50	37	36
less than half	20	54	18	21	56	17	28
almost none	3	48	6	7	53	7	10
How many of your questions require the students to **explain** something...?							
almost all	6	50	16	10	50	12	8
more than half	22	53	29	23	51	26	26
about half	38	51	34	36	51	38	35
less than half	27	50	17	29	54	20	27
almost none	1	39	4	1	42	5	1
How many of your questions require the students to **apply** many things that they know in order to arrive at an answer or solve a problem about something they did not know before...?							
almost all	3	52	9	5	53	11	6
more than half	14	51	20	16	53	21	18
about half	24	49	31	24	49	26	22
less than half	45	53	29	48	51	29	43
almost none	8	49	12	6	49	12	9
ON WRITTEN TESTS:							
How many of your questions require the students to give back only facts and other information they have **memorized**...?							
almost all	14	51	22	10	51	19	10
more than half	33	52	29	31	51	27	24
about half	31	50	30	37	49	30	30
less than half	12	49	15	17	54	17	26

(table continued on next page)

Types of Questions Teachers Ask (continued)

Questionnaire Items	Grade 4	Grade 8			Grade 11		
	Percentage of Teachers	Ave. Test Score	Percentage Stud.	Tchrs.	Ave. Test Score	Percentage Stud.	Tchrs.
How many of your questions require the students to explain something...?							
almost all	1	47	11	2	45	8	3
more than half	13	50	24	10	49	20	16
about half	26	52	32	35	49	34	30
less than half	40	52	26	42	56	27	42
almost none	13	51	8	10	54	10	8
How many of your questions require the students to apply many things that they know in order to arrive at an answer or solve a problem about something they did not know before...?							
almost all	3	50	9	2	55	13	6
more than half	9	49	17	9	52	18	16
about half	14	49	28	20	48	28	17
less than half	43	52	29	50	52	25	42
almost none	25	53	17	16	51	16	17

Table 15

Factors Most Frequently Cited as Serious Problems

Factor	Percent of Teachers Identifying Factor as "Serious Problem" Grade 4	Grade 8	Grade 11
lack of equipment	28	17	20
inadequate budget	23	22	32
lack of student interest in science	7	13	21
mismatch between student reading abilities and science materials	20	17	19
inadequate teacher preparation for teaching science	20	10	7
class size	18	28	20
inadequate articulation of curriculum across grades	15	19	23
lack of public support and recognition of teachers	36	47	58
lack of support from administration	14	14	23
lack of opportunity/support for inservice	19	11	18
lack of assistance from resource person	19	15	18

The last question on the teacher questionnaire at each level was a free-response question asking the teachers to list the two or three areas in which they felt it would be most beneficial to receive inservice training. While many topics were mentioned, the agreement across grade levels on those

most often listed was substantial. Most teachers want ideas -- ideas for laboratory activities and ideas for teaching various topics or concepts. At all grade levels, particularly at the upper levels, teachers want to learn more about the use of computers, both as a tool for science and as an instructional aid. Many teachers, given the opportunity to write their opinions, provided unsolicited views. Generally, the teachers were expressing their concerns about two matters--time to prepare for science instruction, and money (i.e., salaries and funds for supplies). Some commented on the ineffectiveness of inservice and proposed means by which teachers could exchange ideas.

Principal Questionnaire

The principals in the State of Connecticut are 84 to 86 percent male, based on the results from the principal questionnaires administered in schools participating in the assessment. Approximately two-thirds of the principals at the various levels have ten or more years of teaching experience.

The information provided by elementary principals on the number of minutes per week students receive science instruction is consistent with the information provided by the teachers. In approximately one-third of the schools, the fourth graders receive 80 minutes or less of science instruction per week. When asked how science instruction is provided to fourth grade students, 15 percent of the principals indicated "every day," 55 percent indicated "3 or 4 times every week," 18 percent indicated "once or twice every week," and 11 percent indicated "in blocks alternating with other subjects." Ninety-five percent of the junior high principals indicated that the eighth graders have 5 class periods each week in science.

When asked to contrast the emphasis on science now as compared to five years ago, the percentages of principals reporting a recent increase was 56 percent of the elementary principals, 49 percent of the junior high principals, and 68 percent of the high school principals. Almost all the remaining principals indicated that the emphasis has remained constant. Asked if their schools have a teacher who is given released time or additional salary to purchase equipment or help other teachers in science, only 19 percent of the elementary principals reported in the affirmative, compared to 49 percent at the junior high level and 71 percent at the senior high level. Approximately three-quarters of the principals across the three levels reported that they would be willing to sacrifice or have already sacrificed coverage of some traditional science content so as to incorporate units in environmental problems or energy.

The high school principals were asked how much difficulty they have in securing qualified science teachers. Forty-two percent said they have a great deal of difficulty, and another 42 percent indicated that they have some difficulty. These same high school principals generally felt that the crisis in science education (i.e., the widening gap between those who pursue science and the general population) is a problem. Twenty-five percent reported that this crisis was "very serious," and 58 percent reported that it is "somewhat serious."

While most principals across the levels felt that the best orientation for the science programs in their schools is "equal emphasis on academic preparation and scientific literacy" (55 percent to 67 percent at the various

levels), there is a definite difference in the recommended emphases across the grade levels. Academic preparation was defined as the preparation of students for further study or work in science, while scientific literacy was defined in terms of the preparation of all students to deal with science-related problems as individuals or citizens. Most of the elementary principals who did not recommend an equal emphasis felt that more emphasis should be placed on scientific literacy. More senior high principals felt that the emphasis should be placed on academic preparation. Over 90 percent of the senior high principals felt that their science programs were at least adequate for both their college-bound and non-college-bound students; however, 33 percent reported that their programs were outstanding for their college-bound students, while only 13 percent reported that their programs were outstanding for the non-college-bound students.

Like the teachers, the principals were asked how much of a problem each of several factors presents in its effect on science instruction. Generally, very few principals at the three levels reported that these factors were "serious problems". A substantial number (21 percent) of elementary principals reported that inadequate preparation of teachers to teach science was a serious problem. If one looks primarily at the factors the principals considered "somewhat of a problem," the factors that seemed the most problematic to the principals are consistent with those seen as problems by the teachers.

CHAPTER VI

SUMMARY AND RECOMMENDATIONS

The 1984-85 CAEP assessment has provided a great deal of information for science educators in Connecticut to evaluate. The CAEP state advisory committee for science met with members of the State Department of Education and Advanced Systems to review the assessment data and develop the conclusions and recommendations presented in this chapter. It is suggested that the reader examine this report in its entirety, and perhaps request the Technical Report of the assessment to examine the data that provided the basis for this report. Specific findings discussed below are merely selected highlights.

Comparisons with the Nation, the Northeast and Connecticut Five Years Ago

With the exceptions noted below, the performance of Connecticut students on this assessment was similar to the performance of students in the northeast and the nation, and with Connecticut students on the 1979-1980 assessment.

- On questions measuring scientific inquiry, students in grades 4 and 8 scored significantly higher than did students in the northeast and the nation; however, students in grade 11 performed significantly lower than students in Connecticut five years ago and than students in the northeast.

- On questions measuring the life sciences, students in grades 8 and 11 scored significantly lower than their northeast regional counterparts.

- On questions measuring the earth and space sciences, grade 4 students scored significantly higher than students in the northeast and the nation.

- On questions measuring the physical sciences and on the total test, students in Connecticut at grades 4, 8, and 11 performed at about the same levels as their regional and national counterparts and as their Connecticut counterparts on the 1979-80 assessment.

Another way to look at the same data is to view performance by grade level.

- At grade 4, Connecticut students performed significantly better in the areas of scientific inquiry and earth and space science than students in the nation and the northeast region.

- At grade 8, students in Connecticut performed significantly better than their national and northeast counterparts in scientific inquiry and significantly lower than their northeast counterparts in the life sciences.

- At grade 11, the students showed significant declines in scientific inquiry compared to five years ago and they performed significantly lower than their northeast counterparts in scientific inquiry and in the life sciences.

Connecticut Strengths and Weaknesses

Scientific Inquiry

- Fourth grade students were able to measure temperature, time and length with good accuracy. They were able to classify objects according to size, shape, and color. They had greater difficulty understanding factors influencing the speed of falling bodies.

- Students at the eighth and eleventh grade levels were able to measure length accurately, but they were not as accurate with weight measurements and far less able to use these measurements in calculating the density of an object.

- Students in the eighth and eleventh grades were quite successful in constructing a simple electric circuit, but had difficulty reading an ammeter.

- Most eighth and eleventh grade students were able to set up and focus a microscope, but had greater difficulty in drawing an accurate portrayal of what they saw.

- Eighth and eleventh grade students, when presented with a problem involving the design and carrying out of an experiment, generally were able to design an experiment, control for one variable, make the measurements, and record the data. They had difficulty, however, in drawing conclusions relevant to the data gathered and in controlling for more than one variable.

- The reading and interpreting of tables and graphs seemed to be a relative strength at grades 4 and 8. However, drawing conclusions based on data, particularly in response to open-ended questions, appears to be a weakness at all levels.

- Many students at all three grade levels have some misconceptions about the nature of science, feeling that scientists deal only with facts and certainty and must work in laboratories. Also, many students tend to trust authority (written information and the work of scientists) even in situations clearly requiring their own empirical verification or questioning.

- Many students, across the grades, are not precise in making measurements. They either ignore or don't understand directions pertaining to precision (e.g. "nearest centimeter") and have difficulty working with scales graduated in units other than one.

Life Sciences

- Most students at the different grade levels have some knowledge of food groups, but they have difficulty associating particular foods to these groups.

- Grade 11 students seem to know about cells, tissues, etc., but they are not as knowledgeable about the concepts of heredity, genetics, and population stability.

- While many grade 4 students do not understand basic life functions, students understand these functions fairly well by grade 8.

- Students across the grades can identify parts of the body and their general functions, but are weak on the specifics of how they work and interact.

- Few eighth and eleventh grade students could determine the effect of the removal of a link from a simple food chain, or identify probable effects on an ecosystem of the introduction of a new species.

Earth/Space Sciences

- Students at all levels demonstrated a knowledge of facts in astronomy, but they had difficulty in linking ideas. For example, many eighth and eleventh grade students were weak in their understanding of relative motions of bodies, explaining the causes for seasons, and demonstrating an awareness of relative distances in space. Few of these students know that the stars appear to move across the sky at night just as the sun does during the day.

- Similarly, in meteorology, students at all levels can recall many facts. However, grade 8 and 11 students lack an adequate understanding of such things as relative humidity, percent chance of rain, and weather maps--things that students are likely to hear about or see every day.

- Students across all grades have some knowledge of fossils and changes in the earth's surface; few eighth graders, however, could select a reasonable explanation of why fossils of fish might be found on a mountaintop.

- Students at all levels know facts about fuel resources, yet few eleventh graders knew what is wrong with the idea of switching to electricity as a solution to coal, gas, and oil shortages.

Physical Sciences

- Grade 8 and 11 performance was generally poor on test questions dealing with practical concepts of electricity, the understanding of which could be of value for purposes of energy conservation and safety in the home (e.g., energy consumption of heating devices and function of fuses).

- Students at the different levels know many facts about oxidation (e.g., new substances formed, oxygen needed for fire, rusting).

- Relative weaknesses in physics at grades 8 and 11 include a poor understanding of the laws of motion and heat transfer.

- Students across the grades have a poor understanding of the structure of matter (e.g., atoms, elements, and compounds), and few can recognize a situation in which the conservation of matter applies.

Selected Questionnaire Results

- Attitudes toward science became less positive as students porgress through the grades. Sixty-seven percent of the fourth graders, 57 percent of the eighth graders, and 40 percent of the eleventh graders felt that science was either their favorite or one of their favorite courses.

- Grade 4 students reported frequent use of the textbook in class and infrequent involvement in laboratory activity. In the high school, this situation was reversed.

- Many varied teaching techniques used in moderation appear to be associated with higher student achievement than nonuse or overuse of a particular technique.

- Most grade 8 and 11 students are assigned regular science homework, and over 40 percent of the students spend one half to one hour per day doing that homework.

- The course-taking patterns of grade 11 students indicate that most students (91 percent) have taken, are taking, or intend to take biology. The corresponding percentages for chemistry and physics are considerably lower (62 percent and 41 percent respectively).

- Most students in all three grades felt that science will be useful to them in every day life.

- Students in all three grades tested reported that they very much enjoyed doing science experiments in groups (81 percent in grade 4, 78 percent in grade 8, and 66 percent in grade 11). In all grades, students who enjoyed these activities performed significantly better on the assessment than those who did not.

Teacher Questionnaire

- Grade 4 teachers were asked how many minutes per week their students receive of science instruction. In many schools, relatively little time is spent on science instruction at grade 4.

Amount of Science Instruction per Week	Percent of Grade 4 Teachers Selecting Response
40 minutes or less	9
41 - 80 minutes	27
81 - 120 minutes	35
121 - 160 minutes	17
more than 160 minutes	7

- According to teachers, the most valuable resources for information on new developments in education are other teachers and professional publications.

- The teachers reported that the questions they ask aloud in class and on written tests tend to require the recall of facts more than the explanation of phenomena or application of concepts.

- Twenty-six percent of the fourth grade teachers had an undergraduate college major in an academic field; the remainder of the respondents majored in education. For the eighth grade science teachers sampled, the number of academic majors increased to 62 percent, while for science teachers in grades 9-12 the nubmer increased to 84 percent.

- Five percent of the fourth grade teachers had an undergraduate major in science. This number increased to 52 percent in the eighth grade, and to 77 percent for teachers in grades 9-12.

- Fourth grade teachers attend workshops more frequently than they take college courses in science or in the teaching of science. A greater percentage of teachers in the eighth and eleventh grades take college courses and attend conferences than do those in the fourth grade.

- Additional comments returned with teachers' questionnaires indicate that they are greatly concerned about the availability of time for classroom preparation and the amount of money for science materials.

Principal Questionnaire

- Almost half of the principals at grades 4 and 11 felt that the lack of materials posed a problem for their science programs while 30 percent of the grade 8 principals felt that way.

- Less than 40 percent of the eleventh grade principals felt that the introduction of a computer to their schools had resulted in increased student achievement. This percentage decreased in both the eighth and fourth grades.

- Elementary school principals expressed more concern about the lower status of science compared to other subjects than did principals at the secondary levels.

- Seventy-nine percent of the elementary schools responded that they do not have a teacher who is given released time or extra pay to help others with their science teaching or with the purchase of materials.

- One-half to two-thirds of the principals across the levels reported that the emphasis on science in their schools had increased in the past five years.

- High school principals reported that although they generally believe that their science programs were adequate, they felt that they were better suited for college-bound than for non-college-bound students. Eighty-three percent of the high school principals felt that the

widening gap in science literacy between those who pursue science as a career and the general population was a problem.

- Seventy-one percent of the elementary school principals felt that inadequate teacher preparation posed a problem for their science programs.

- High school principals generally reported having difficulty securing qualified science teachers to fill vacancies.

- Principals generally had positive views about science curriculum development activities undertaken in their schools.

Recommendations

The recommendations below, based on the review of assessment data by the science advisory committee, the State Department of Education, and the contractor for the assessment, pertain to the teaching of science, the evaluations, development, and modification of science programs, and administrative support for science teachers and programs.

Recommendation 1

Science programs at all levels should place greater emphasis on hands-on laboratory activities. Laboratory experiences can be a strong motivator, and "doing" science is generally an effective method for increasing understanding and retention.

Recommendation 2

Teachers of science at all levels should give greater attention to higher order thinking processes in science. Both the state and local school systems should improve and increase inservice activities providing training in the teaching of higher order thinking skills in science.

Recommendation 3

Science teachers at all levels should employ a variety of methods and materials to improve students' attitudes toward science and their perceptions of the utility of science.

Recommendation 4

Real-life, personal and societal applications of scientific principles and concepts should be emphasized in science programs. Current topics and issues in science should also be a part of these programs.

Recommendation 5

More time must be devoted to science instruction in the elementary grades. Instructional time in science should at least meet the minimum recommendation of 135 minutes per week as stated in the State Department of Education's Guide to Curriculum Development in Science.

Recommendation 6

Every elementary school should have a science resource teacher who can assist other teachers in selecting, securing, updating, and utilizing science instructional methods, materials, and equipment.

Recommendation 7

The high school program for many students should include more science study, particularly in courses in the physical sciences.

Recommendation 8

Science teachers and curriculum developers should strive to achieve a better balance of instruction in the life, earth, and physical sciences for all students.

Recommendation 9

High school teachers and curriculum developers should evaluate science courses for students not bound for college to be certain that they include the content and skills that are required to function successfully within a technological society.

Recommendation 10

Elementary school inservice programs should provide teachers with a strong background in science content and with a variety of instructional techniques for use in the classroom.

Recommendation 11

Teachers should be given time not only to prepare classes, but to develop and update curricula and to participate in inservice activities.

Recommendation 12

All science classrooms should be well supplied with up-to-date materials and equipment.

Recommendation 13

Given the lack of emphasis on science instruction in the elementary grades, the state should develop a set of science concepts, skills, and evaluation instruments appropriate for use in the elementary schools.

APPENDIX

SAMPLE TEST

SAMPLE TEST

As noted in the preface, a sample test is included as an appendix to this report to encourage teachers to make the results more relevant to their own students. Teachers may wish to copy this sample test and administer it to students in their classes. Then they can compute the percentage of correct responses achieved by their students on each item, compare it to the statewide results, and determine whether the area measured by the item is a possible area of weakness for their students. To assist in the interpretation of results, the statewide results are given below in addition to comments about the items and references to pages in this report where further discussion of the topics covered by the items can be found.

Item	Key	Percent Answering Correctly Statewide Grade 8	Grade 11	Comments	Page Reference in Report
1	A	58	73	misconceptions about science and scientists	10
2	C	67	--	experimental/control groups; sampling	11,12
3	D	48	60	misconception about exactness of science; poor knowledge of measurement error	14
4	B	51	67	poor interpolation skills	17
5	B	--	31	general, but not in-depth understanding of human body	19
6	B	24	32	little knowledge of photosynthesis	20
7	C	--	67	disappointing grade 11 performance	20,21
8	B	19	24	media-induced misconception leading to preference for option A	21
9	B	52	--	little concept of relative distances in space	22
10	E	31	40	poor understanding of relative humidity, despite frequent use of term in weather reports	24
11	C	29	--	poor understanding of formation and changing of earth's surface	25
12	A	--	38	difficulty with higher order thinking--applying knowledge of energy	25
13	B	56	57	grade 4 did better by 15%	26,27
14	A	50	65	practical knowledge in the home (generally weak in electricity)	28
15	B	28	50	little understanding of types of matter	28,29
16	C	51	49	poor performance on simple application of conservation of matter	28,29

1. Which of the following statements about science or scientists is generally true?

   A. The solution a scientist finds to one problem often raises new questions or problems.

   B. Scientists deal with facts and fixed rules and not with assumptions and uncertainty.

   C. Scientists know all there is to know about many areas such as cell reproduction and composition of the earth.

   D. True scientific experiments must be conducted in laboratories for one to be sure of the results.

2. Sarah wants to compare the effects of breakfast cereal and lab food on the growth of mice. She could BEST do this by

   A. feeding a mouse cereal and recording its growth.

   B. feeding a mouse standard lab food and recording its growth; then switching it to a diet of cereal and recording its growth.

   C. finding two groups of identical mice; feeding one cereal and one standard lab food and recording their growth.

   D. feeding several mice standard lab food; then adding cereal to their diet to see if their growth rate changes.

3. Whenever scientists carefully measure any quantity many times, they expect that

   A. all of the measurements will be exactly the same.

   B. only two of the measurements will be exactly the same.

   C. all but one of the measurements will be exactly the same.

   D. most of the measurements will be close but not exactly the same.

   E. I don't know.

AUTOMOBILE STOPPING DISTANCES

Speed (kilometers per hour)	Thinking Distance (meters)	Braking Distance (meters)	Total Stopping Distance (meters)
40	8	10	18
60	12	23	35
80	17	42	59
100	21	64	85

4. Which of the following would be the best estimate of the TOTAL STOPPING DISTANCE for a car traveling at 66 kilometers per hour?

   A. 30 meters

   B. 40 meters

   C. 50 meters

   D. 60 meters

5. Which is the primary function of the veins in the human body?

   A. carry blood from the heart to other parts of the body

   B. carry blood to the heart from other parts of the body

   C. carry blood both away from and to the heart

   D. allow the exchange of materials between the blood and body cells

   E. I don't know

6. What are two products produced by plants during photosynthesis?

   A. sugar and carbon dioxide

   B. sugar and oxygen

   C. water and carbon dioxide

   D. water and oxygen

   E. sugar and water

7. In a particular meadow there are many rabbits that eat the grass. There are also many hawks that eat the rabbits. Last year a disease broke out among the rabbits and a great number of them died. Which of the following probably then occurred?

   A. The grass died and the hawk population decreased.
   B. The grass died and the hawk population increased.
   C. The grass grew taller and the hawk population decreased.
   D. The grass grew taller and the hawk population increased.
   E. Neither the grass nor the hawks were affected by the death of the rabbits.
   F. I don't know.

8. If water used to cool nuclear reactors causes harm when it is returned to the environment, the harm is most likely to occur because of the water's

   A. radioactivity.
   B. high temperature.
   C. density.
   D. chemical sludge.

9. Which one of the following BEST explains why telescopes used for studying the stars are often built on mountain tops?

   A. to get as close as possible to the stars
   B. to get above layers of smog, dust and fog
   C. to make sure that the powerful telescopes can be placed on a firm platform
   D. to get away from the noise of large cities
   E. I don't know.

10. The statement that the relative humidity is 50 percent means that

   A. the chance of rain is 50 percent.
   B. the atmosphere contains 50 pounds of water per cubic mile.
   C. the atmosphere contains 50 grams of water per cubic meter.
   D. the atmosphere would be saturated with water if the air temperature were 50°F.
   E. the atmosphere contains half as much water as it could contain at its present temperature.
   F. I don't know.

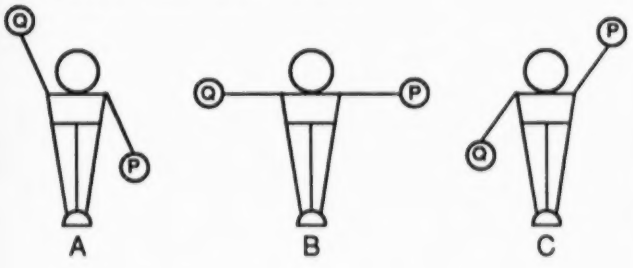
11. A fossil of an ocean fish was found in a rock outcrop on a mountain. This probably means that

   A. fish once lived on the mountain.
   B. the relative humidity was once very high.
   C. the mountain was raised up after the fish died.
   D. fish used to be amphibians like toads and frogs.
   E. the fossil fish was probably carried to the mountain by a great flood.
   F. I don't know.

12. Some people think that the solution to coal, gas and oil shortages is to switch over to electricity. What is wrong with this idea?

   A. Most electricity is produced from coal, gas and oil.
   B. If we switch over to electricity, many people will lose their jobs.
   C. It is more expensive to build machines and vehicles that run on electricity.
   D. Electricity is far too expensive.
   E. There is nothing wrong with this idea.
   F. I don't know.

13. Suppose that you want to drop a penny and a quarter at exactly the same time and have them hit the floor at exactly the same time. Which picture BEST shows how you would hold the penny and the quarter just before you drop them?



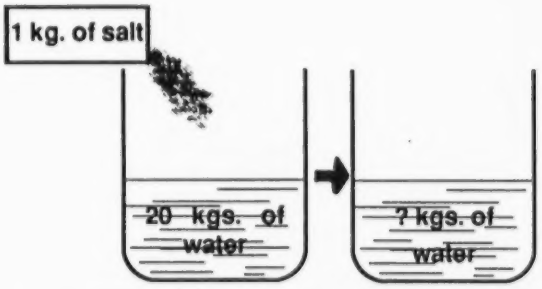
D. I don't know.

14. Which of these appliances consumes the most energy in 15 minutes of continuous operation?

A. electric clothes dryer
B. color television
C. vacuum cleaner
D. electric can opener

15. Water is

A. an element.
B. a compound.
C. a solution.
D. a mixture.



16. One kilogram of salt is completely dissolved in twenty kilograms of water. The resulting saltwater will weigh

A. nineteen kilograms.
B. twenty kilograms.
C. twenty-one kilograms.
D. The weight is unpredictable.
E. I don't know.

Connecticut State Department of Education

Program and Support Services

Lorraine M. Aronson
Deputy Commissioner

Office of Research and Evaluation

Pascal D. Forgione, Jr., Chief

Joan B. Baron
CAEP Program Director

Division of Curriculum and Professional Development

Betty J. Sternberg, Director

Sigmund Abeles, Consultant,
Science

Velma A. Adams, Editor